

## NUCLEOSIDIC AND NON-NUCLEOSIDIC FOLATE CONJUGATES

### CROSS-REFERENCE

This application is a continuation-in-part application of U.S. Application Serial No. 09/098,166, filed June 16,  
5 1998.

### FIELD OF THE INVENTION

The present invention is directed to mononucleosides and oligonucleotides that are conjugated to folic acid, related folates, antifolates and analogs thereof. The  
10 present invention further provides non-nucleosidic folate conjugates. The present invention also provides methods for preparation of nucleosidic and non-nucleosidic conjugates.

### BACKGROUND OF THE INVENTION

Protein synthesis is directed by nucleic acids through  
15 the intermediacy of messenger RNA (mRNA). Antisense methodology is the complementary hybridization of relatively short oligonucleotides to mRNA or DNA such that the normal, essential functions, such as protein synthesis, of these intracellular nucleic acids are disrupted. Hybridization is  
20 the sequence-specific hydrogen bonding via Watson-Crick base pairs of oligonucleotides to RNA or single-stranded DNA. Such base pairs are said to be complementary to one another.

The naturally occurring events that provide the

disruption of the nucleic acid function, discussed by Cohen in *Oligonucleotides: Antisense Inhibitors of Gene Expression*, CRC Press, Inc., Boca Raton, Fl (1989) are thought to be of two types. The first, hybridization arrest, describes the terminating event in which the oligonucleotide inhibitor binds to the target nucleic acid and thus prevents, by simple steric hindrance, the binding of essential proteins, most often ribosomes, to the nucleic acid. Methyl phosphonate oligonucleotides: Miller, P.S. and Ts'O, P.O.P. (1987) *Anti-Cancer Drug Design*, 2:117-128, and  $\alpha$ -anomer oligonucleotides are the two most extensively studied antisense agents which are thought to disrupt nucleic acid function by hybridization arrest.

Another means by which antisense oligonucleotides disrupt nucleic acid function is by hybridization to a target mRNA, followed by enzymatic cleavage of the targeted RNA by intracellular RNase H. A 2'-deoxyribofuranosyl oligonucleotide or oligonucleotide analog hybridizes with the targeted RNA and this duplex activates the RNase H enzyme to cleave the RNA strand, thus destroying the normal function of the RNA. Phosphorothioate oligonucleotides are the most prominent example of an antisense agent that operates by this type of antisense terminating event.

Considerable research is being directed to the application of oligonucleotides and oligonucleotide analogs as antisense agents for diagnostics, research applications and potential therapeutic purposes. At least for therapeutic purposes, the antisense oligonucleotides and oligonucleotide analogs must be transported across cell membranes or taken up by cells to exhibit their activity. However, due to the large size and unfavorable charge-size ratio of oligonucleotides, their cellular uptake is very limited. Numerous efforts have focused on increasing this membrane permeability and cellular delivery of oligonucleotides.

Efforts aimed at improving the transmembrane delivery of nucleic acids and oligonucleotides have utilized protein carriers, antibody carriers, liposomal delivery systems, electroporation, direct injection, cell fusion, viral  
5 vectors, and calcium phosphate mediated transformation. However, many of these techniques are limited by the types of cells in which transmembrane transport is enabled and by the conditions needed for achieving such transport. An alternative that is particularly attractive for the  
10 transmembrane delivery of oligonucleotides is the modification of the physico-chemical properties of oligonucleotides via conjugation to a molecule that facilitates transport.

One method for increasing membrane or cellular  
15 transport of oligonucleotides is the attachment of a pendant lipophilic group. Ramirez, F., Mandal, S.B. and Marecek, J.F., (1982) *J. Am. Chem. Soc.*, 104:5483, introduced the phospholipid group 5'-O-(1,2-di-O-myristoyl-sn-glycero-3-phosphoryl) into the dimer TpT independently at the 3' and  
20 5' positions. Subsequently Shea, R.G., Marsters, J.C. and Bischofberger, N. (1990), *Nuc. Acids Res.*, 18:3777, disclosed oligonucleotides having a 1,2-di-O-hexyldecyl-rac-glycerol group linked to a 5'-phosphate on the 5'-terminus of the oligonucleotide. Certain of the Shea et. al. authors  
25 also disclosed these and other compounds in patent application PCT/US90/01002. A further glucosyl phospholipid was disclosed by Guerra, F.I., Neumann, J.M. and Hynh-Dinh, T. (1987), *Tetrahedron Letters*, 28:3581.

In other work, a cholesteryl group was attached to the  
30 inter-nucleotide linkage between the first and second nucleotides (from the 3' terminus) of an oligonucleotide. This work is disclosed in United States patent 4,958,013 and further in Letsinger, R.L., Zhang, G., Sun, D.K., Ikeuchi, T. and Sarin, P.S. (1989), *Proc. Natl. Acad. Sci. USA*,  
35 86:6553. Additional approaches to the delivery and study

of oligonucleotides have involved the conjugation of a variety of other molecules and reporter groups. The aromatic intercalating agent anthraquinone was attached to the 2' position of a sugar fragment of an oligonucleotide as reported by Yamana, K., Nishijima, Y., Ikeda, T., Gokota, T. Ozaki, H., Nakano, H., Sangen, O. and Shimidze, T. (1990) *Bioconjugate Chem.*, 1:319; Lemairte, M., Bayard, B. and Lebleu, B. (1986), *Proc. Natl. Acad. Sci. USA*, 84:648; and Leonetti, J.-P., Degols, G. and Lebleu, B. (1990), *Bioconjugate Chem.*, 1:149. Lysine and polylysines have also been conjugated to oligonucleotides to improve their charge-size characteristics. The poly(L-lysine) was linked to the oligonucleotide via periodate oxidation of the 3'-terminal ribose followed by reduction and coupling through a N-morpholine ring. Oligonucleotide-poly(L-lysine) conjugates are described in European Patent application 87109348.0. In this instance the lysine residue was coupled to a 5' or 3' phosphate of the 5' or 3' terminal nucleotide of the oligonucleotide. A disulfide linkage has also been utilized at the 3' terminus of an oligonucleotide to link a peptide to the oligonucleotide as is described by Corey, D.R. and Schultz, P.G. (1987), *Science*, 238:1401; Zuckermann, R.N., Corey, D.R., and Schultz, P.G. (1988), *J. Am. Chem. Soc.*, 110:1614; and Corey, D.R., Pei, D. and Schultz, P.G. (1989), *J. Am. Chem. Soc.*, 111:8524.

Nelson, P.S., Frye, R.A. and Liu, E. (1989), *Nuc. Acids Res.*, 17:7187 describe a linking reagent for attaching biotin to the 3'-terminus of an oligonucleotide. This reagent, N-Fmoc-O-DMT-3-amino-1,2-propanediol is now commercially available from Clontech Laboratories (Palo Alto, CA) under the name 3'-Amine on. It is also commercially available under the name 3'-Amino-Modifier reagent from Glen Research Corporation (Sterling, VA). This reagent was also utilized to link a peptide to an oligonucleotide as reported by Judy, C.D., Richardson, C.D.

and Brousseau, R. (1991), *Tetrahedron Letters*, 32:879. A similar commercial reagent (actually a series of such linkers having various lengths of polymethylene connectors) for linking to the 5'-terminus of an oligonucleotide is 5'-

5 Amino-Modifier C6. These reagents are available from Glen Research Corporation (Sterling, VA). These compounds or similar ones were utilized by Krieg, A.M., Gmelig-Meyling, F., Gourley, M.F., Kisch, W.J., Chrisey, L.A. and Steinberg, A.D. (1991), *Antisense Research and Development*, 1:161 to

10 link fluorescein to the 5'-terminus of an oligonucleotide. Other compounds of interest have also been linked to the 3'-terminus of an oligonucleotide. Asseline, U., Delaure, M., Lancelot, G., Toulme, F., Thuong, N.T., Montenay-Garestier, T. and Helene, C. (1984), *Proc. Natl. Acad. Sci. USA*,

15 81:3297 described linking acridine on the 3'-terminal phosphate group of an poly (Tp) oligonucleotide via a polymethylene linkage. Haralambidis, J., Duncan, L. and Tregear, G.W. (1987), *Tetrahedron Letters*, 28:5199 report building a peptide on a solid state support and then linking

20 an oligonucleotide to that peptide via the 3' hydroxyl group of the 3' terminal nucleotide of the oligonucleotide. Chollet, A. (1990), *Nucleosides & Nucleotides*, 9:957 attached an Aminolink 2 (Applied Biosystems, Foster City, CA) to the 5' terminal phosphate of an oligonucleotide.

25 They then used the bifunctional linking group SMPB (Pierce Chemical Co., Rockford, Il) to link an interleukin protein to the oligonucleotide.

Conjugation of lipids, reporters, peptides and other molecules to oligonucleotides is not limited to the terminal

30 3' and 5'-positions. A wide variety of conjugates have also been reported in the literature wherein attachment is performed at any one or more of the 2'-positions on the nucleotide building blocks of the oligonucleotide. Further conjugates have also been reported wherein attachment occurs

35 on the internucleotide linkage or on one of the atoms of the

nucleobase of any one of the nucleotide units of the oligonucleotide. For example, an EDTA iron complex has been linked to the 5 position of a pyrimidine nucleoside as reported by Dreyer, G.B. and Dervan, P.B. (1985), *Proc.*

- 5 *Natl. Acad. Sci. USA*, 82:968. Fluorescein has been linked to an oligonucleotide in the same manner as reported by Haralambidis, J., Chai, M. and Tregear, G.W. (1987), *Nucleic Acid Research*, 15:4857 and biotin in the same manner as described in PCT application PCT/US/02198. Fluorescein,  
10 biotin and pyrene were also linked in the same manner as reported by Telser, J., Cruickshank, K.A., Morrisison, L.E. and Netzel, T.L. (1989), *J. Am. Chem. Soc.*, 111:6966. A commercial reagent, Amino-Modifier-dT, from Glen Research Corporation (Sterling, VA) can be utilized to introduce  
15 pyrimidine nucleotides bearing similar linking groups into oligonucleotides.

- Manoharan et al. PCT Application WO 93/07883, have also reported the conjugation of oligonucleotides with a variety of molecules such as steroids, reporter molecules,  
20 reporter enzymes, vitamins, non-aromatic lipophilic molecules, chelators, porphyrins, intercalators, peptides and proteins through the intermediacy of varied linking groups, such as w-aminoalkoxy and w-aminoalkylamino groups. Conjugation has been reported at the 3'-, 5'-, 2'-,  
25 internucleotide linkage and nucleobase positions of oligonucleotides. Such oligonucleotide conjugates are expected to have improved physicochemical properties that facilitated their uptake and delivery into cells as demonstrated by *in vitro* experiments.

- 30 The intracellular and intranuclear delivery of nucleic acids and oligonucleotides, however, is still a challenge. Most often, penetration of heretofore reported oligonucleotide conjugates has been found to be limited. This has typically been a problem because such conjugates  
35 have generally been designed to improve the passive

absorption of the oligonucleotides where the size, physico-chemical properties and extra-cellular concentration of the conjugate play important limiting roles. This coupled with the limited extra-cellular stability of nucleic acids and oligonucleotides demands the development of novel conjugates that will deliver higher levels of nucleic acids and oligonucleotides into specific tissues and targeted cells. One such approach for delivering oligonucleotides, is to exploit the active transport mechanism of receptor mediated endocytosis.

Unlike many of the methods mentioned above, receptor mediated endocytotic activity can be used successfully both *in vitro* and *in vivo*. This mechanism of uptake involves the movement of ligands bound to membrane receptors into the interior of an area that is enveloped by the membrane via invagination of the membrane structure. This process is initiated via activation of a cell-surface or membrane receptor following binding of a specific ligand to the receptor. Many receptor-mediated endocytotic systems are known and have been studied, including those that recognize sugars such as galactose, mannose, mannose-6-phosphate, peptides and proteins such as transferrin, asialoglycoprotein, vitamin B12, insulin and epidermal growth factor (EGF). Receptor mediated endocytosis has been well studied and is known to be a critical pathway for the uptake and internalization of a variety of cellular nutrients. These are highly developed mechanisms because of their critical role in providing nutrients to cells and in maintaining cellular physiology. Thus many examples of the utilization of receptor mediated endocytosis pathways for the delivery of drugs, proteins, nucleic acids and other molecules to cells are known. One way in which this has been applied is the conjugation of essential nutrients that are actively transported into cells with the drug or molecule of interest. The transporters or receptors

involved in the uptake are capable of recognizing the nutrient portion of the conjugate and ferrying the entire conjugate into the cell. Examples of nutrients that are actively transported into cells and that may be of use in conjugates include, but are not limited to, folic acid, vitamin B6, cholesterol and vitamin B12. Such molecules have been conjugated to macromolecules such as nucleic acids and oligonucleotides to afford conjugates that exhibit improved cellular penetration. Manorharan et al., PCT Application WO 93/07883; Low et al., U.S. Patent 5,108,921, U.S. Patent 5,416,016.

Many vitamins possess an acid or alcohol functionality that is readily modified for conjugation to oligonucleotides. For example, conjugation of an N-hydroxy succinimide ester of an acid moiety of retinoic acid to an amine function on a linker pendant to an oligonucleotide resulted in a oligonucleotide-Vitamin A conjugate attached via an amide bond. Retinol has been converted to its phosphoramidite and conjugated to the 5'-terminus of oligonucleotides via a phosphodiester or phosphorothioate linkage. Likewise, vitamin E and vitamin B6 may also be conjugated to oligonucleotides to improve transport into cells.

Pyridoxal (vitamin B<sub>6</sub>) has specific B<sub>6</sub>-binding proteins. The role of these proteins in pyridoxal transport has been studied by Zhang and McCormick, *Proc. Natl. Acad. Sci. USA*, **1991** 88, 10407. Zhang and McCormick also have shown that a series of N-(4'-pyridoxyl)amines, in which several synthetic amines were conjugated at the 4'-position of pyridoxal, are able to enter cells by a process facilitated by the B6 transporter. They also demonstrated the release of these synthetic amines within the cell. Other pyridoxal family members include pyridoxine, pyridoxamine, pyridoxal phosphate, and pyridoxic acid. Pyridoxic acid, niacin, pantothenic acid, biotin, folic acid



and ascorbic acid can be conjugated to oligonucleotides using N-hydroxysuccinimide esters that are reactive with aminolinkers located on the oligonucleotide, as described above for retinoic acid.

5 Folic acid and its various forms, such as dihydrofolate and tetrahydrofolate, are essential vitamins that are crucial for the biosynthesis of nucleic acids and therefore are critical to the survival and proliferation of cells. Folate cofactors play an important role in the one-  
10 carbon transfers that are critical for the biosynthesis of pyrimidine nucleosides. Cells therefore have a sophisticated system of transporting folates into the cytoplasm. Uptake of folates occurs by two different pathways depending on the cell type. Cells expressing a  
15 carrier or transporter for folate that exhibits a low affinity ( $K_d \sim 10^{-6}$  M) for the vitamin prefer reduced over oxidized forms of folate. Cells that express membrane receptors called folate binding protein (FBP), in contrast, exhibit high binding affinity ( $K_d \sim 10^{-9}$  M) and prefer the  
20 oxidized form of the vitamin. This latter receptor is believed to mediate the uptake of folates into the cytoplasm via endocytosis.

The use of biotin conjugates and also folic acid conjugates to enhance transmembrane transport of exogenous  
25 molecules, including oligonucleotides, has been reported by Low et al., U.S. Patent 5,108,921; U.S. Patent 5,416,016; PCT Application WO 90/12096. Folic acid was conjugated to 3'-aminoalkyl-oligonucleotides at their 3'-terminus via carbodiimide chemistry. The multiplicity of folate  
30 receptors on membrane surfaces of most cells and the associated receptor mediated endocytotic processes were implicated in the enhanced transport of these oligonucleotide-folic acid conjugates into cells. There are however, several limitations to this approach for the  
35 conjugation of folic acid to oligonucleotides.

Folic acid and many related folates and antifolates exhibit very poor solubility that hinders the effective conjugation of folic acid to oligonucleotides and subsequent purification of oligonucleotide-folic acid conjugates.

5 Further folic acid bears two reactive carboxylic acid groups that are just as likely to react with the terminal amino group of the 3-aminoalkyl-oligonucleotide. Thus conjugation will typically result in a mixture of  $\alpha$ - and  $\gamma$ - conjugates arising from the reaction of the  $\alpha$ -carboxylate and the  $\gamma$ -carboxylate of the glutamic acid portion of the folic acid molecule. This poses difficulties from the standpoint of characterizing the conjugate and further from the standpoint of polyglutamylation of folates. Polyglutamylation of folates is a well recognized phenomenon that has significant implications on the transport, localization and activity of folates. Since polyglutamylation rates differ significantly between the  $\alpha$ - and  $\gamma$ -carboxylates, the use of poorly defined mixtures of oligonucleotide-folate conjugates, as obtained from the Low et al. procedure, U.S. Patent 5,108,921, will lead to variable transport and concentration of the conjugate. Further, the conjugation of folates onto one end of an oligonucleotide may be a disadvantage because of the known propensity of exonucleases to rapidly cleave oligonucleotides by excising the terminal residues. Also, it has been observed that oligonucleotide-folic acid conjugates prepared in this fashion are light sensitive.

Therefore, there is a clear need for new oligonucleotide-folate conjugates, and methods for their preparation, that address the shortcomings of oligonucleotide conjugates as described above. The present invention is directed to this very important end.

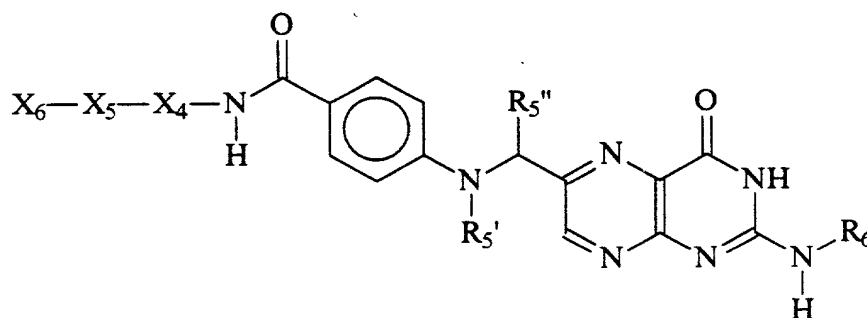
#### SUMMARY OF THE INVENTION

The present invention provides nucleosidic and non-nucleosidic folate conjugates. The present invention also

provides methods for preparation of such conjugates bound to folic acid and related folates. The present invention also provides oligonucleotide-folate conjugates.

The present invention further provides novel  
 5 approaches to the facile synthesis of nucleoside- and oligonucleotide-folate conjugates via a step-wise construction of the folate conjugate. The methods of the present invention are broadly applicable to the attachment  
 10 of a wide variety of folates at any one of many possible sites of attachment on oligonucleotides.

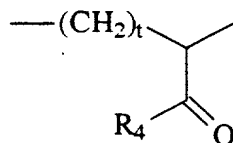
The present invention provides compounds of formula XVII:



XVII

wherein:

$X_4$  is  $-\text{CH}(X_4)$  or a group of formula:



$X_4$  is the side chain of a naturally-occurring or non-naturally-occurring amino acid, or a protected side chain of  
 20 a naturally-occurring or non-naturally-occurring amino acid;

$X_5$  is  $-\text{N}(X_6)\text{C}(=\text{O})-$ ,  $-\text{C}(=\text{O})\text{NH}-$ ,  $-\text{NHC}(=\text{O})-$ ,  $-\text{OC}(=\text{O})\text{NH}-$ ,  
 $-\text{C}(=\text{S})\text{NH}-$ ,  $-\text{SC}(=\text{S})\text{NH}-$ ,  $-\text{SC}(=\text{O})\text{NH}-$ ,  $-\text{OC}(=\text{S})\text{NH}-$ ,  $-\text{C}(=\text{O})\text{O}-$ ,  
 $-\text{C}(=\text{O})(\text{CH}_2)_n-$  or a bond;

$n$  is an integer from 1 to 50;

25 each  $X_6$  and  $X_6$  is, independently, hydrogen or a

hydrocarbyl group selected from  $C_1$ - $C_{10}$  alkyl,  $C_2$ - $C_{10}$  alkenyl,  $C_2$ - $C_{20}$  alkynyl,  $C_6$ - $C_{14}$  aryl,  $C_6$ - $C_{14}$  aralkyl,  $C_3$ - $C_{14}$  cycloalkyl,  $C_5$ - $C_{14}$  fused cycloalkyl,  $C_4$ - $C_{14}$  heterocycle,  $C_4$ - $C_{14}$  heterocyclalkyl,  $C_4$ - $C_{14}$  heteroaryl and  $C_4$ - $C_{14}$

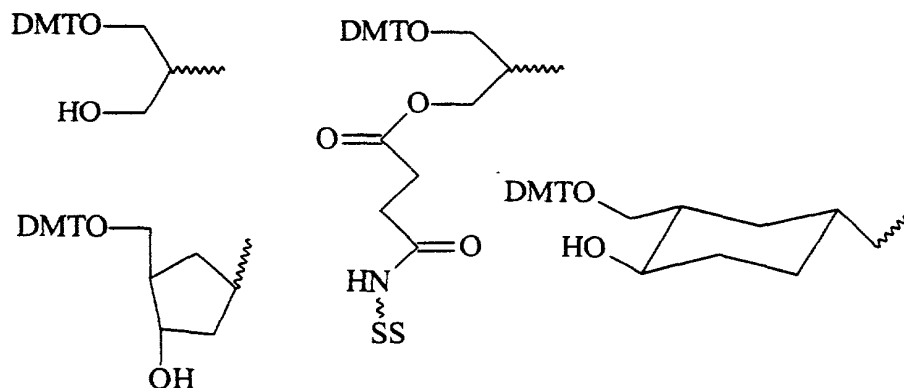
- 5 heteroarylalkyl; wherein said hydrocarbyl group is substituted with at least two hydroxyl groups, and is optionally substituted with oxo, acyl, alkoxy, alkoxycarbonyl, alkyl, alkenyl, alkynyl, amino, amido, azido, aryl, heteroaryl, carboxylic acid, cyano, guanidino, halo, haloalkyl, haloalkoxy, hydrazino, ODMT, alkylsulfonyl, nitro, sulfide, disulfide, sulfone, sulfonate, sulfonamide, thiol, and thioalkoxy; provided that  $X_6$  is not hydrogen;

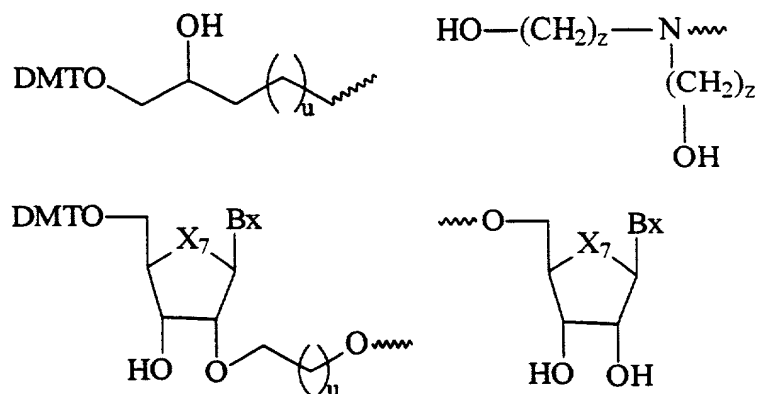
$R_4$  is a hydroxyl group or a protected hydroxyl group;

- 15  $R_5$  is hydrogen,  $C_1$ - $C_{10}$  alkyl,  $C_2$ - $C_{10}$  alkenyl,  $C_2$ - $C_{20}$  alkynyl,  $C_6$ - $C_{14}$  aryl or an amino protecting group;  $R_5$  is hydrogen,  $C_1$ - $C_{10}$  alkyl,  $C_2$ - $C_{10}$  alkenyl,  $C_2$ - $C_{20}$  alkynyl,  $C_6$ - $C_{14}$  aryl,  $C_6$ - $C_{14}$  aralkyl,  $C_3$ - $C_{14}$  cycloalkyl,  $C_5$ - $C_{14}$  fused cycloalkyl, formyl, aminoalkyl or hydroxymethyl;  $R_6$  is hydrogen or an amino-protecting group; and  $t$  is 1 or 2.

20 In a preferred embodiment, the  $X_4$  is aminocaproic acid.

In a preferred embodiment,  $X_6$  has one of the formulas:





wherein:

SS is a solid support;

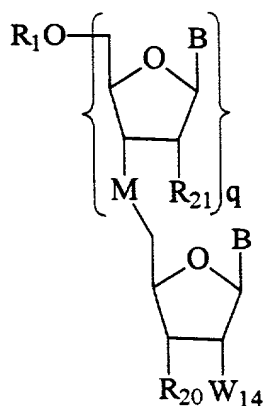
X<sub>7</sub> is O or CH<sub>2</sub>;

Bx is a nucleobase, C<sub>4</sub>-C<sub>14</sub> heterocyclyl or hydrogen;

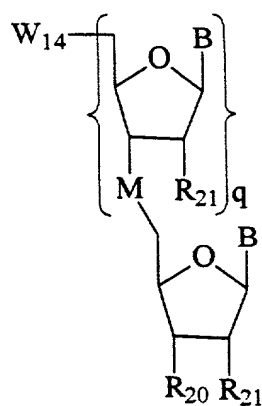
z is an integer from 1 to 50; and

u is an integer from 2 to 5.

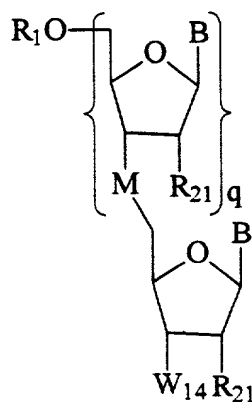
The present invention also provides compounds having formula XVIIA, XVIIIB, XVIC or XVID:



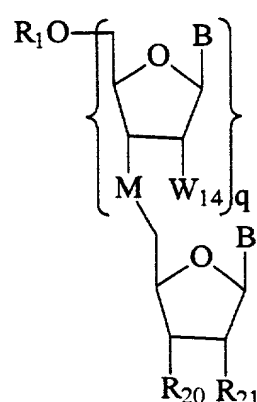
XVIA



XVIC



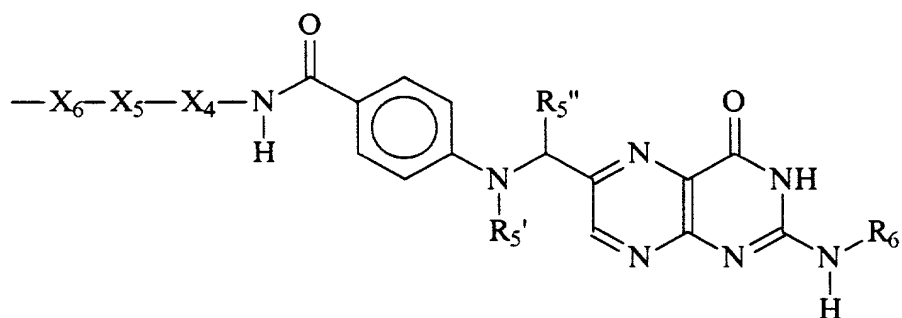
XVIB



XVID

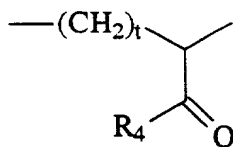
wherein:

$W_{14}$  has the formula:



wherein:

5  $X_4$  is  $-CH(X_4)$  or a group of formula:



$\text{R}_4$  is the side chain of a naturally-occurring or non-naturally-occurring amino acid, or a protected side chain of a naturally-occurring non-naturally-occurring amino acid;

$t$  is 1 or 2;

- 5  $\text{X}_5$  is  $-\text{N}(\text{X}_6)\text{C}(\text{O})-$ ,  $-\text{C}(\text{O})\text{NH}-$ ,  $-\text{NHC}(\text{O})-$ ,  $-\text{OC}(\text{O})\text{NH}-$ ,  $-\text{C}(\text{S})\text{NH}-$ ,  $-\text{SC}(\text{S})\text{NH}-$ ,  $-\text{SC}(\text{O})\text{NH}-$ ,  $-\text{OC}(\text{S})\text{NH}-$ ,  $-\text{C}(\text{O})\text{O}-$ ,  $-\text{C}(\text{O})(\text{CH}_2)_n-$  or a bond;

$n$  is an integer from 1 to 50;

- each  $\text{X}_6$  and  $\text{X}_6'$  is, independently, a bond, hydrogen or a  
 10 hydrocarbyl group selected from  $\text{C}_1\text{-C}_{10}$  alkyl,  $\text{C}_2\text{-C}_{10}$  alkenyl,  $\text{C}_2\text{-C}_{20}$  alkynyl,  $\text{C}_6\text{-C}_{14}$  aryl,  $\text{C}_6\text{-C}_{14}$  aralkyl,  $\text{C}_3\text{-C}_{14}$  cycloalkyl,  $\text{C}_5\text{-C}_{14}$  fused cycloalkyl,  $\text{C}_4\text{-C}_{14}$  heterocycle,  $\text{C}_4\text{-C}_{14}$  heterocyclalkyl,  $\text{C}_4\text{-C}_{14}$  heteroaryl and  $\text{C}_4\text{-C}_{14}$  heteroarylalkyl; wherein said hydrocarbyl group is  
 15 substituted with at least two hydroxyl groups, and is optionally substituted with oxo, acyl, alkoxy, alkoxy carbonyl, alkyl, alkenyl, alkynyl, amino, amido, azido, aryl, heteroaryl, carboxylic acid, cyano, guanidino, halo, haloalkyl, haloalkoxy, hydrazino, ODMT, alkylsulfonyl,  
 20 nitro, sulfide, disulfide, sulfone, sulfonate, sulfonamide, thiol, and thioalkoxy; provided that  $\text{X}_6$  is not hydrogen and  $\text{X}_6'$  is not a bond;

$\text{R}_1$  is hydrogen or a hydroxyl protecting group;

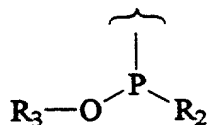
$\text{R}_4$  is a hydroxyl group or a protected hydroxyl group;

- 25 each  $\text{R}_5$  and  $\text{R}_{10}$  is, independently, hydrogen,  $\text{C}_1\text{-C}_{10}$  alkyl,  $\text{C}_2\text{-C}_{10}$  alkenyl,  $\text{C}_2\text{-C}_{20}$  alkynyl,  $\text{C}_6\text{-C}_{14}$  aryl or an amino protecting group;

- $\text{R}_5'$  is hydrogen,  $\text{C}_1\text{-C}_{10}$  alkyl,  $\text{C}_2\text{-C}_{10}$  alkenyl,  $\text{C}_2\text{-C}_{20}$  alkynyl,  $\text{C}_6\text{-C}_{14}$  aryl,  $\text{C}_6\text{-C}_{14}$  aralkyl,  $\text{C}_3\text{-C}_{14}$  cycloalkyl,  $\text{C}_5\text{-C}_{14}$  fused cycloalkyl, formyl, aminoalkyl or hydroxymethyl;  
 30

R<sub>6</sub> is hydrogen or an amino-protecting group;

R<sub>20</sub> is hydrogen or a group of Formula:



R<sub>2</sub> is -N(R<sub>7</sub>)<sub>2</sub>, or a heterocycloalkyl or  
5 heterocycloalkenyl ring containing from 4 to 7 atoms, and  
having up to 3 heteroatoms selected from nitrogen, sulfur,  
and oxygen;

R<sub>7</sub> is straight or branched chain alkyl having from 1 to  
10 carbons;

10 R<sub>3</sub> is a phosphorus protecting group;

R<sub>21</sub> is hydrogen, hydroxyl, fluoro or a group of formula  
Z-R<sub>22</sub>-(R<sub>23</sub>)<sub>v</sub>;

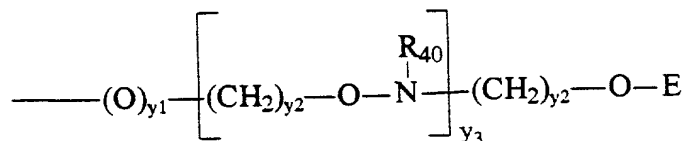
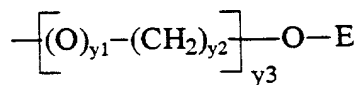
Z is O, S, NH or N-R<sub>22</sub>-(R<sub>23</sub>)<sub>v</sub>;

R<sub>22</sub> is C<sub>1</sub>-C<sub>20</sub> alkyl, C<sub>2</sub>-C<sub>20</sub> alkenyl, or C<sub>2</sub>-C<sub>20</sub> alkynyl;

15 R<sub>23</sub> is hydrogen, amino, halogen, hydroxyl, thiol, keto,  
carboxyl, nitro, nitroso, nitrile, trifluoromethyl,  
trifluoromethoxy, O-alkyl, S-alkyl, NH-alkyl, N-dialkyl, O-  
aryl, S-aryl, NH-aryl, O-aralkyl, S-aralkyl, NH-aralkyl,  
amino, N-phthalimido, imidazole, azido, hydrazino,  
20 hydroxylamino, isocyanato, sulfoxide, sulfone, sulfide,  
disulfide, silyl, aryl, heterocycle, carbocycle, inter-  
calator, reporter molecule, conjugate, polyamine, polyamide,  
polyalkylene glycol, polyether, a group that enhances the  
pharmacodynamic properties of oligonucleotides, or a group  
25 that enhances the pharmacokinetic properties of oligonucleo-  
tides;

or R<sub>21</sub> has one of the formulas:





wherein:

y<sub>1</sub> is 0 or 1;

each y<sub>2</sub> is, independently, 0 to 10;

5 y<sub>3</sub> is 1 to 10;

E is N(R<sub>41</sub>)(R<sub>42</sub>) or N=C(R<sub>41</sub>)(R<sub>42</sub>);

each R<sub>41</sub> and each R<sub>42</sub> is independently H, C<sub>1</sub>-C<sub>10</sub> alkyl, a nitrogen protecting group, or R<sub>41</sub> and R<sub>42</sub> taken together form a nitrogen protecting group; or R<sub>41</sub> and R<sub>42</sub> taken together with the N or C atom to which they are attached form a ring structure that can include at least one heterotom selected from N and O;

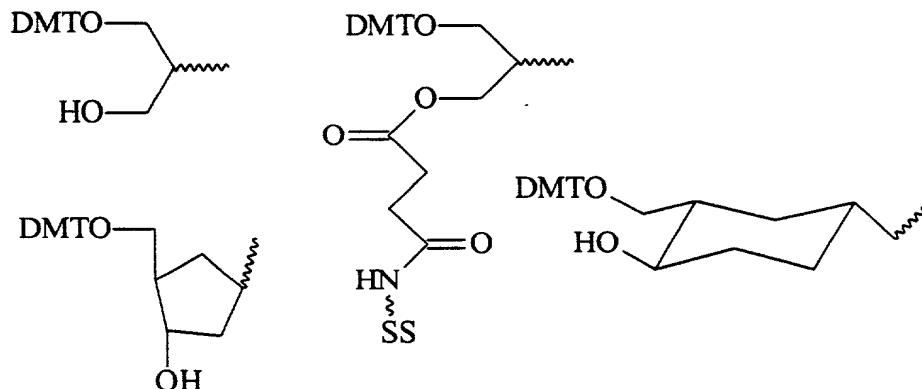
B is a nucleobase;

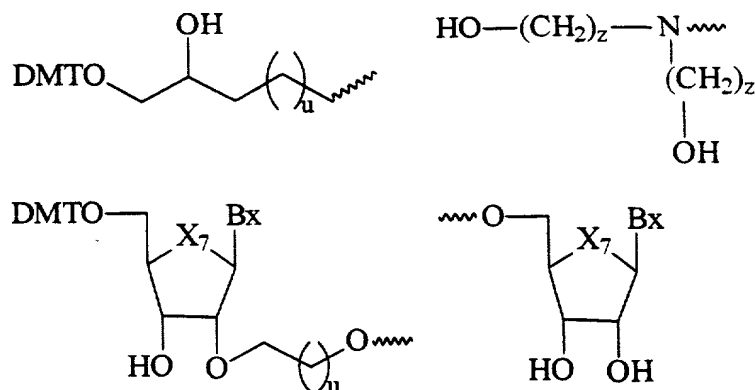
M is an optionally protected internucleoside linkage;

15 q is from zero to about 50; and

v is from zero to about 10.

In a preferred embodiment of the present invention, X<sub>6</sub> has one of formulas:





wherein:

SS is a solid support;

X<sub>7</sub> is O or CH<sub>2</sub>;

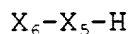
Bx is a nucleobase, C<sub>4</sub>-C<sub>14</sub> heterocyclyl or hydrogen;

5 z is an integer from 1 to 50; and

u is an integer from 2 to 5.

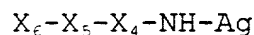
The present invention also provides methods for the preparation of compounds of formula XVII comprising the steps of:

10 (a) providing a hydroxy compound of formula:



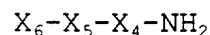
(b) protecting said hydroxyl groups of X<sub>6</sub> with protecting groups to form a protected hydroxy compound;

(c) reacting said protected hydroxy compound with an amino-protected amino acid to form a covalently linked hydroxy compound of formula:



wherein Ag is an amino protecting group; and

(d) cleaving the amino-protecting group of said covalently linked hydroxy compound to form a hydroxy compound bearing a deprotected amino group and having formula:



(e) reacting said amino group with a folate moiety;

and

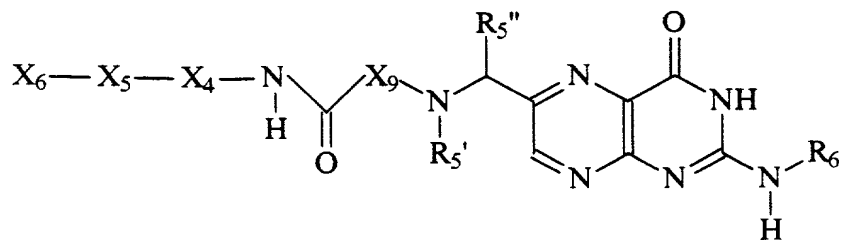
T06007 T060200

(f) cleaving the protecting groups on said hydroxyl groups of step (b) to form a compound of formula XVII.

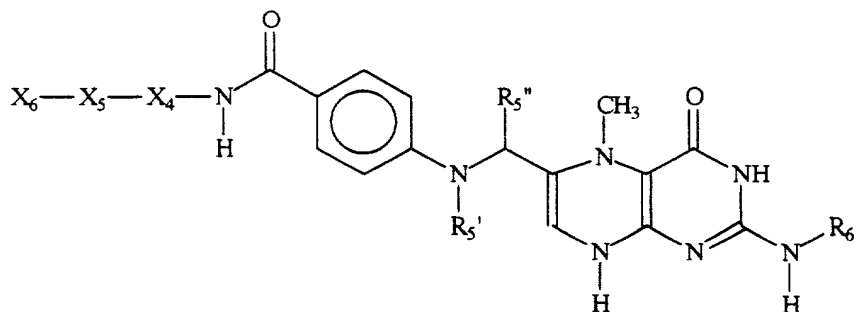
The methods of the present invention also provide compounds of the formula XVII wherein R<sub>4</sub> is a hydroxyl group protected with C<sub>1</sub>-C<sub>10</sub> alkyl, C<sub>2</sub>-C<sub>10</sub> alkenyl or C<sub>2</sub>-C<sub>20</sub> alkynyl, the methods further comprising the step of deprotecting the R<sub>4</sub> group with a deprotecting reagent. In a preferred embodiment, R<sub>4</sub> is a methoxy group or an allyl protected hydroxyl group.

10 In one preferred embodiment, the deprotection reagent is an aqueous amine. In another preferred embodiment, the deprotection reagent is a combination of aqueous amine and a mercapto compound.

In one aspect of the present invention, the folate moiety of step (e) is selected such that a compound of the formula shown below is synthesized:



In another aspect of the present invention, the folate moiety of step (e) is selected such that a compound of the formula shown below is synthesized:

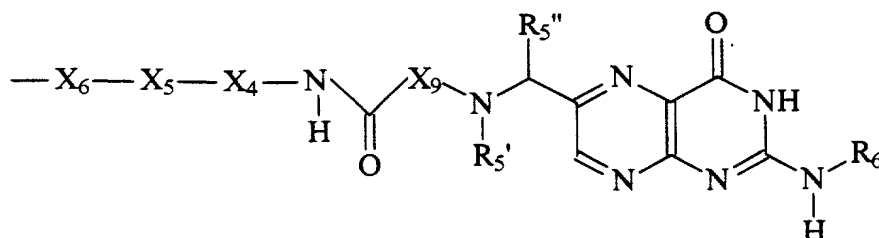


The present invention also provides methods further comprising the steps of:

(g) protecting one of said hydroxyl groups of  $X_6$  with a dimethoxytrityl group; and

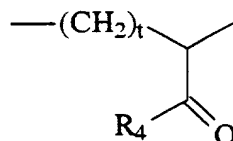
5 (h) phosphitylating the other of said hydroxyl groups of  $X_6$ .

The present invention also provides compounds of the formula:



wherein:

10  $X_4$  is  $-CH(X_4')$  or a group of formula:



$X_4'$  is the side chain of a naturally-occurring or non-naturally-occurring amino acid, or a protected side chain of a naturally-occurring or non-naturally-occurring amino acid;

15  $X_5$  is  $-N(X_6')C(O)-$ ,  $-C(O)NH-$ ,  $-NHC(O)-$ ,  $-OC(O)NH-$ ,  $-C(S)NH-$ ,  $-SC(S)NH-$ ,  $-SC(O)NH-$ ,  $-OC(S)NH-$ ,  $-C(O)O-$ ,  $-C(O)(CH_2)_n-$  or a bond;

$n$  is an integer from 1 to 50;

20 each  $X_6$ ,  $X_6'$  and  $X_9$  is, independently, a bond, hydrogen or a hydrocarbyl group selected from  $C_1-C_{10}$  alkyl,  $C_2-C_{10}$  alkenyl,  $C_2-C_{20}$  alkynyl,  $C_6-C_{14}$  aryl,  $C_6-C_{14}$  aralkyl,  $C_3-C_{14}$  cycloalkyl,  $C_5-C_{14}$  fused cycloalkyl,  $C_4-C_{14}$  heterocycle,  $C_4-C_{14}$  heterocyclalkyl,  $C_4-C_{14}$  heteroaryl and  $C_4-C_{14}$

heteroarylalkyl; wherein said hydrocarbyl group is substituted with at least two hydroxyl groups, and is optionally substituted with oxo, acyl, alkoxy, alkoxy carbonyl, alkyl, alkenyl, alkynyl, amino, amido, azido, aryl, heteroaryl, carboxylic acid, cyano, guanidino, halo, haloalkyl, haloalkoxy, hydrazino, ODMT, alkylsulfonyl, nitro, sulfide, disulfide, sulfone, sulfonate, sulfonamide, thiol, and thioalkoxy; provided that each of  $X_6$  and  $X_6$  is not a bond and  $X_6$  is not hydrogen;

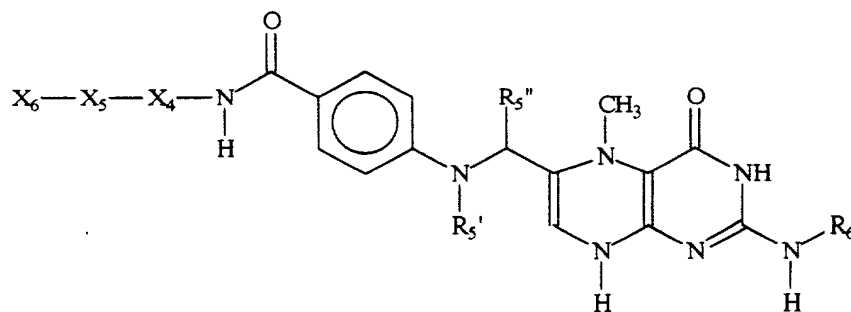
10  $R_4$  is a hydroxyl group or a protected hydroxyl group;

$R_5$  is hydrogen,  $C_1$ - $C_{10}$  alkyl,  $C_2$ - $C_{10}$  alkenyl,  $C_2$ - $C_{20}$  alkynyl,  $C_6$ - $C_{14}$  aryl or an amino-protecting group

$R_{5'}$  is hydrogen,  $C_1$ - $C_{10}$  alkyl,  $C_2$ - $C_{10}$  alkenyl,  $C_2$ - $C_{20}$  alkynyl,  $C_6$ - $C_{14}$  aryl,  $C_6$ - $C_{14}$  aralkyl,  $C_3$ - $C_{14}$  cycloalkyl, formyl, 15 aminoalkyl or hydroxymethyl;

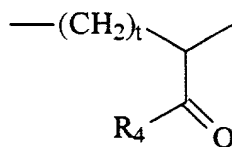
$R_6$  is hydrogen or an amino-protecting group; and  $t$  is 1 or 2.

The present invention further provides compounds of the formula:



20 wherein:

$X_4$  is  $-CH(X_4')$  or a group of formula:



$X_4'$  is the side chain of a naturally-occurring or non-naturally-occurring amino acid, or a protected side chain of

a naturally-occurring or non-naturally-occurring amino acid;

$X_5$  is  $-N(X_6)C(O)-$ ,  $-C(O)NH-$ ,  $-NHC(O)-$ ,  $-OC(O)NH-$ ,  
 $-C(S)NH-$ ,  $-SC(S)NH-$ ,  $-SC(O)NH-$ ,  $-OC(S)NH-$ ,  $-C(O)O-$ ,  
 $-C(O)(CH_2)_n-$  or a bond;

5         $n$  is an integer from 1 to 50;

each  $X_6$  and  $X_6'$  is, independently, hydrogen or a  
hydrocarbyl group selected from  $C_1$ - $C_{10}$  alkyl,  $C_2$ - $C_{10}$  alkenyl,  
 $C_2$ - $C_{20}$  alkynyl,  $C_6$ - $C_{14}$  aryl,  $C_6$ - $C_{14}$  aralkyl,  $C_3$ - $C_{14}$  cycloalkyl,  
 $C_5$ - $C_{14}$  fused cycloalkyl,  $C_4$ - $C_{14}$  heterocycle,  $C_4$ - $C_{14}$   
10 heterocyclylalkyl,  $C_4$ - $C_{14}$  heteroaryl and  $C_4$ - $C_{14}$   
heteroarylalkyl; wherein said hydrocarbyl group is  
substituted with at least two hydroxyl groups, and is  
optionally substituted with oxo, acyl, alkoxy,  
alkoxycarbonyl, alkyl, alkenyl, alkynyl, amino, amido, azido,  
15 aryl, heteroaryl, carboxylic acid, cyano, guanidino, halo,  
haloalkyl, haloalkoxy, hydrazino, ODMT, alkylsulfonyl, nitro,  
sulfide, disulfide, sulfone, sulfonate, sulfonamide, thiol,  
and thioalkoxy; provided that  $X_6$  is not hydrogen;

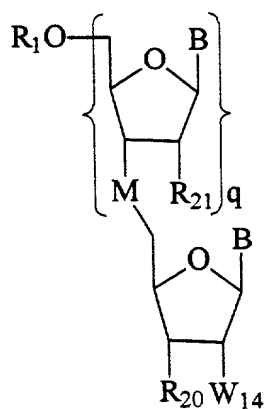
$R_4$  is a hydroxyl group or a protected hydroxyl group;

20         $R_5$  is hydrogen,  $C_1$ - $C_{10}$  alkyl,  $C_2$ - $C_{10}$  alkenyl,  $C_2$ - $C_{20}$   
alkynyl,  $C_6$ - $C_{14}$  aryl or an amino-protecting group

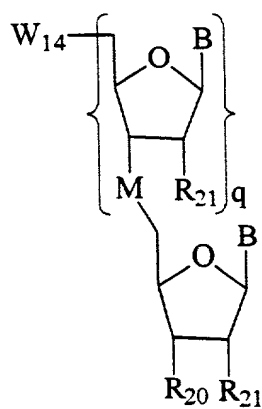
$R_5'$  is hydrogen,  $C_1$ - $C_{10}$  alkyl,  $C_2$ - $C_{10}$  alkenyl,  $C_2$ - $C_{20}$   
alkynyl,  $C_6$ - $C_{14}$  aryl,  $C_6$ - $C_{14}$  aralkyl,  $C_3$ - $C_{14}$  cycloalkyl, formyl,  
aminoalkyl or hydroxymethyl;

25         $R_6$  is hydrogen or an amino protecting group; and  
 $t$  is 1 or 2.

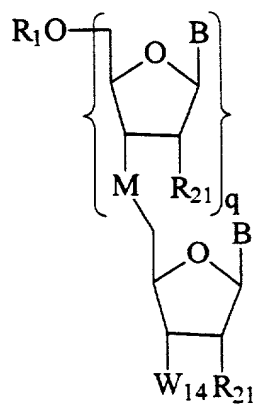
The present invention also provides compounds of the  
formula XVIA, XVIB, XVIC or XVID:



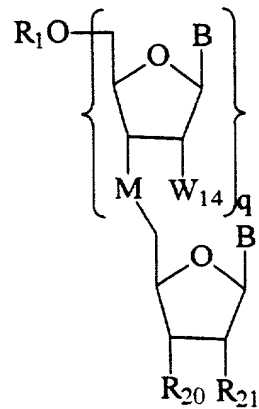
XVI A



XVI C



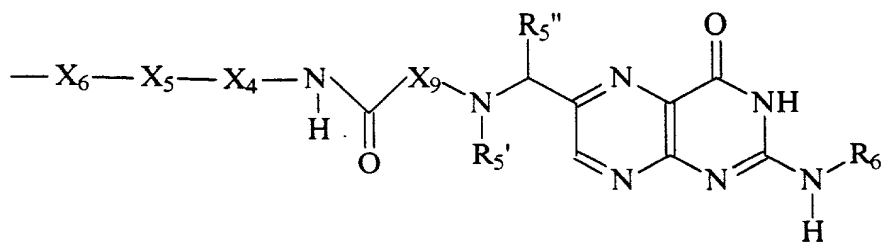
XVI B



XVI D

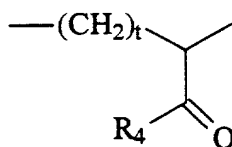
wherein:

$W_{14}$  has the formula:



wherein:

$X_4$  is  $-\text{CH}(X_4)$  or a group of formula:



$X_4$  is the side chain of a naturally-occurring or non-naturally-occurring amino acid, or a protected side chain of a naturally-occurring or non-naturally-occurring amino acid;

$t$  is 1 or 2;

$X_5$  is  $-\text{N}(X_6)\text{C}(\text{O})-$ ,  $-\text{C}(\text{O})\text{NH}-$ ,  $-\text{NHC}(\text{O})-$ ,  $-\text{OC}(\text{O})\text{NH}-$ ,  $-\text{C}(\text{S})\text{NH}-$ ,  $-\text{SC}(\text{S})\text{NH}-$ ,  $-\text{SC}(\text{O})\text{NH}-$ ,  $-\text{OC}(\text{S})\text{NH}-$ ,  $-\text{C}(\text{O})\text{O}-$ ,  $-\text{C}(\text{O})(\text{CH}_2)_n-$  or a bond;

$n$  is an integer from 1 to 50;

each  $X_6$ ,  $X_6$ , and  $X_9$  is, independently, a bond, hydrogen or a hydrocarbyl group selected from  $\text{C}_1$ - $\text{C}_{10}$  alkyl,  $\text{C}_2$ - $\text{C}_{10}$  alkenyl,  $\text{C}_2$ - $\text{C}_{20}$  alkynyl,  $\text{C}_6$ - $\text{C}_{14}$  aryl,  $\text{C}_6$ - $\text{C}_{14}$  aralkyl,  $\text{C}_3$ - $\text{C}_{14}$  cycloalkyl,  $\text{C}_5$ - $\text{C}_{14}$  fused cycloalkyl,  $\text{C}_4$ - $\text{C}_{14}$  heterocycle,  $\text{C}_4$ - $\text{C}_{14}$  heterocyclalkyl,  $\text{C}_4$ - $\text{C}_{14}$  heteroaryl and  $\text{C}_4$ - $\text{C}_{14}$  heteroarylalkyl; wherein said hydrocarbyl group is substituted with at least two hydroxyl groups, and is optionally substituted with oxo, acyl, alkoxy, alkoxycarbonyl, alkyl, alkenyl, alkynyl, amino, amido, azido, aryl, heteroaryl, carboxylic acid, cyano, guanidino, halo, haloalkyl, haloalkoxy, hydrazino, ODMT, alkylsulfonyl, nitro, sulfide, disulfide, sulfone, sulfonate, sulfonamide, thiol, and thioalkoxy; provided that each  $X_6$  and  $X_9$  is not hydrogen and  $X_6$  is not a bond;

$\text{R}_1$  is hydrogen or a hydroxyl protecting group;

$\text{R}_4$  is a hydroxyl group or a protected hydroxyl group;

each  $\text{R}_5$  and  $\text{R}_{40}$  is, independently, hydrogen,  $\text{C}_1$ - $\text{C}_{10}$  alkyl,  $\text{C}_2$ - $\text{C}_{10}$  alkenyl,  $\text{C}_2$ - $\text{C}_{20}$  alkynyl,  $\text{C}_6$ - $\text{C}_{14}$  aryl or an amino-protecting group

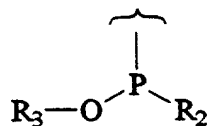
$\text{R}_{50}$  is hydrogen,  $\text{C}_1$ - $\text{C}_{10}$  alkyl,  $\text{C}_2$ - $\text{C}_{10}$  alkenyl,  $\text{C}_2$ - $\text{C}_{20}$  alkynyl,  $\text{C}_6$ - $\text{C}_{14}$  aryl,  $\text{C}_6$ - $\text{C}_{14}$  aralkyl,  $\text{C}_3$ - $\text{C}_{14}$  cycloalkyl, formyl,



aminoalkyl or hydroxymethyl;

$R_6$  is hydrogen or an amino protecting group;

$R_{20}$  is hydrogen or a group of formula:



5  $R_2$  is  $-N(R_7)_2$ , or a heterocycloalkyl or heterocycloalkenyl ring containing from 4 to 7 atoms, and having up to 3 heteroatoms selected from nitrogen, sulfur, and oxygen;

10  $R_7$  is straight or branched chain alkyl having from 1 to 10 carbons;

$R_3$  is a phosphorus protecting group;

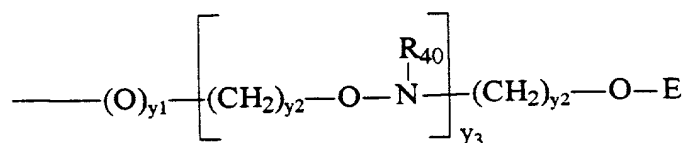
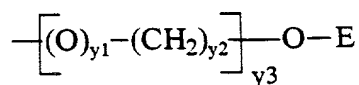
$R_{21}$  is hydrogen, hydroxyl, fluoro or a group of formula  $Z-R_{22}-(R_{23})_v$ ;

$Z$  is O, S, NH or  $N-R_{22}-(R_{23})_v$ ;

15  $R_{22}$  is  $C_1-C_{20}$  alkyl,  $C_2-C_{20}$  alkenyl, or  $C_2-C_{20}$  alkynyl;

20  $R_{23}$  is hydrogen, amino, halogen, hydroxyl, thiol, keto, carboxyl, nitro, nitroso, nitrile, trifluoromethyl, trifluoromethoxy, O-alkyl, S-alkyl, NH-alkyl, N-dialkyl, O-aryl, S-aryl, NH-aryl, O-aralkyl, S-aralkyl, NH-aralkyl, amino, N-phthalimido, imidazole, azido, hydrazino, hydroxylamino, isocyanato, sulfoxide, sulfone, sulfide, disulfide, silyl, aryl, heterocycle, carbocycle, intercalator, reporter molecule, conjugate, polyamine, polyamide, polyalkylene glycol, polyether, a group that enhances the  
25 pharmacodynamic properties of oligonucleotides, or a group that enhances the pharmacokinetic properties of oligonucleotides;

or  $R_{21}$  has one of the formulas:



wherein:

y<sub>1</sub> is 0 or 1;

each y<sub>2</sub> is, independently, 0 to 10;

5 y<sub>3</sub> is 1 to 10;

E is N(R<sub>41</sub>)(R<sub>42</sub>) or N=C(R<sub>41</sub>)(R<sub>42</sub>);

each R<sub>41</sub> and each R<sub>42</sub> is independently H, C<sub>1</sub>-C<sub>10</sub> alkyl, a nitrogen protecting group, or R<sub>41</sub> and R<sub>42</sub> taken together form a nitrogen protecting group; or R<sub>41</sub> and R<sub>42</sub> taken together with  
10 the N or C atom to which they are attached form a ring structure that can include at least one heterotom selected from N and O;

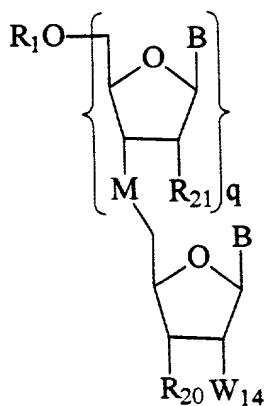
B is a nucleobase;

M is an optionally protected internucleoside linkage;

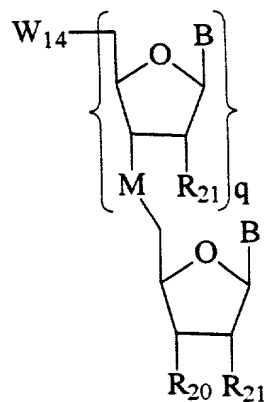
15 q is from zero to about 50; and

v is from zero to about 10.

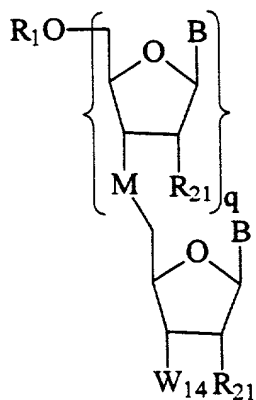
The present invention further provides compounds of the formula XVIA, XVIB, XVIC or XVID:



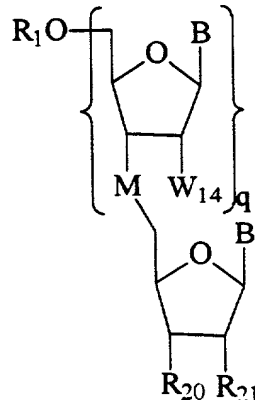
XVIA



XVIC



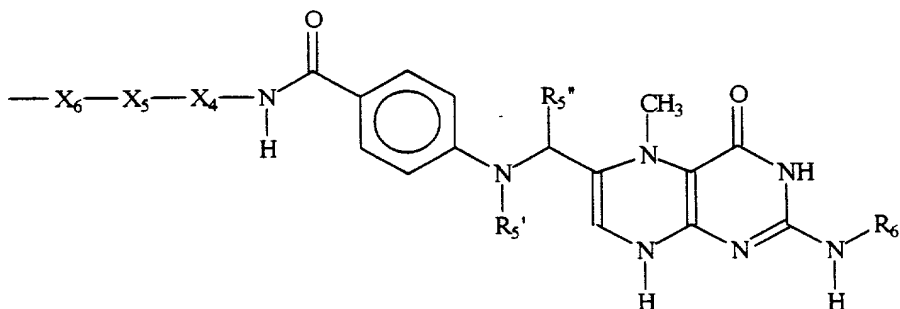
XVIB



XVID

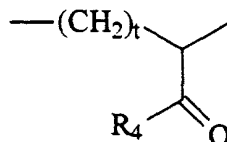
wherein:

5      W<sub>14</sub> has the formula:



wherein:

$X_4$  is  $-\text{CH}(X_4')$  or a group of formula:



$X_4$  is the side chain of a naturally-occurring or non-naturally-occurring amino acid, or a protected side chain of a naturally-occurring or non-naturally-occurring amino acid;

$t$  is 1 or 2;

$X_5$  is  $-\text{N}(X_6')\text{C}(\text{O})-$ ,  $-\text{C}(\text{O})\text{NH}-$ ,  $-\text{NHC}(\text{O})-$ ,  $-\text{OC}(\text{O})\text{NH}-$ ,  $-\text{C}(\text{S})\text{NH}-$ ,  $-\text{SC}(\text{S})\text{NH}-$ ,  $-\text{SC}(\text{O})\text{NH}-$ ,  $-\text{OC}(\text{S})\text{NH}-$ ,  $-\text{C}(\text{O})\text{O}-$ ,  
10  $-\text{C}(\text{O})(\text{CH}_2)_n-$  or a bond;

$n$  is an integer from 1 to 50;

each  $X_6$  and  $X_6'$  is, independently, a bond, hydrogen or a hydrocarbyl group selected from  $\text{C}_1$ - $\text{C}_{10}$  alkyl,  $\text{C}_2$ - $\text{C}_{10}$  alkenyl,  $\text{C}_2$ - $\text{C}_{20}$  alkynyl,  $\text{C}_6$ - $\text{C}_{14}$  aryl,  $\text{C}_6$ - $\text{C}_{14}$  aralkyl,  $\text{C}_3$ - $\text{C}_{14}$  cycloalkyl,  
15  $\text{C}_5$ - $\text{C}_{14}$  fused cycloalkyl,  $\text{C}_4$ - $\text{C}_{14}$  heterocycle,  $\text{C}_4$ - $\text{C}_{14}$  heterocyclalkyl,  $\text{C}_4$ - $\text{C}_{14}$  heteroaryl and  $\text{C}_4$ - $\text{C}_{14}$  heteroarylalkyl; wherein said hydrocarbyl group is substituted with at least two hydroxyl groups, and is optionally substituted with oxo, acyl, alkoxy,  
20 alkoxy carbonyl, alkyl, alkenyl, alkynyl, amino, amido, azido, aryl, heteroaryl, carboxylic acid, cyano, guanidino, halo, haloalkyl, haloalkoxy, hydrazino, ODMT, alkylsulfonyl, nitro, sulfide, disulfide, sulfone, sulfonate, sulfonamide, thiol, and thioalkoxy; provided that  $X_6$  is not hydrogen and  $X_6'$  is  
25 not a bond;

$\text{R}_1$  is hydrogen or a hydroxyl protecting group;

$\text{R}_4$  is a hydroxyl group or a protected hydroxyl group;

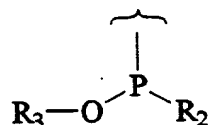
each  $\text{R}_5$  and  $\text{R}_{40}$  is, independently, hydrogen,  $\text{C}_1$ - $\text{C}_{10}$  alkyl,  $\text{C}_2$ - $\text{C}_{10}$  alkenyl,  $\text{C}_2$ - $\text{C}_{20}$  alkynyl,  $\text{C}_6$ - $\text{C}_{14}$  aryl or an amino-  
30 protecting group

$\text{R}_{50}$  is hydrogen,  $\text{C}_1$ - $\text{C}_{10}$  alkyl,  $\text{C}_2$ - $\text{C}_{10}$  alkenyl,  $\text{C}_2$ - $\text{C}_{20}$  alkynyl,  $\text{C}_6$ - $\text{C}_{14}$  aryl,  $\text{C}_6$ - $\text{C}_{14}$  aralkyl,  $\text{C}_3$ - $\text{C}_{14}$  cycloalkyl, formyl,

aminoalkyl or hydroxymethyl;

$R_6$  is hydrogen or an amino protecting group;

$R_{20}$  is hydrogen or a group of formula:



5  $R_2$  is  $-N(R_7)_2$ , or a heterocycloalkyl or heterocycloalkenyl ring containing from 4 to 7 atoms, and having up to 3 heteroatoms selected from nitrogen, sulfur, and oxygen;

$R_7$  is straight or branched chain alkyl having from 1 to 10 10 carbons;

$R_3$  is a phosphorus protecting group;

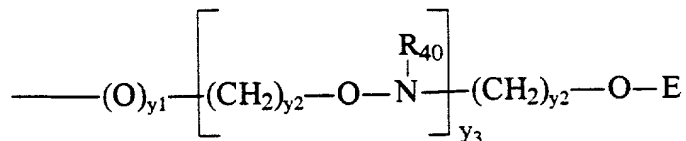
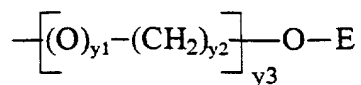
$R_{21}$  is hydrogen, hydroxyl, fluoro or a group of formula  $Z-R_{22}-(R_{23})_v$ ;

$Z$  is O, S, NH or  $N-R_{22}-(R_{23})_v$ ;

15  $R_{22}$  is  $C_1-C_{20}$  alkyl,  $C_2-C_{20}$  alkenyl, or  $C_2-C_{20}$  alkynyl;

$R_{23}$  is hydrogen, amino, halogen, hydroxyl, thiol, keto, carboxyl, nitro, nitroso, nitrile, trifluoromethyl, trifluoromethoxy, O-alkyl, S-alkyl, NH-alkyl, N-dialkyl, O-aryl, S-aryl, NH-aryl, O-aralkyl, S-aralkyl, NH-aralkyl, 20 amino, N-phthalimido, imidazole, azido, hydrazino, hydroxylamino, isocyanato, sulfoxide, sulfone, sulfide, disulfide, silyl, aryl, heterocycle, carbocycle, intercalator, reporter molecule, conjugate, polyamine, polyamide, polyalkylene glycol, polyether, a group that enhances the 25 pharmacodynamic properties of oligonucleotides, or a group that enhances the pharmacokinetic properties of oligonucleotides;

or  $R_{21}$  has one of the formulas:



wherein:

y<sub>1</sub> is 0 or 1;

each y<sub>2</sub> is, independently, 0 to 10;

5 y<sub>3</sub> is 1 to 10;

E is N(R<sub>41</sub>)(R<sub>42</sub>) or N=C(R<sub>41</sub>)(R<sub>42</sub>);

each R<sub>41</sub> and each R<sub>42</sub> is independently H, C<sub>1</sub>-C<sub>10</sub> alkyl, a  
nitrogen protecting group, or R<sub>41</sub> and R<sub>42</sub> taken together form a  
nitrogen protecting group; or R<sub>41</sub> and R<sub>42</sub> taken together with  
10 the N or C atom to which they are attached form a ring  
structure that can include at least one heterotom selected  
from N and O;

B is a nucleobase;

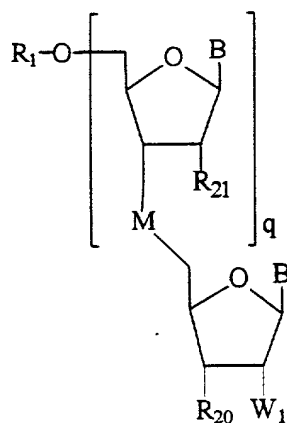
M is an optionally protected internucleoside linkage;

15 q is from zero to about 50; and

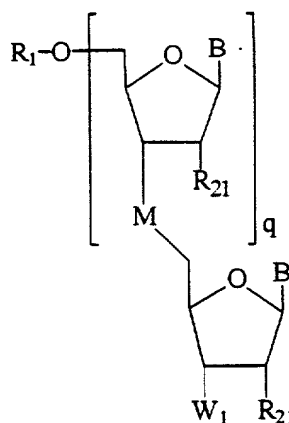
v is from zero to about 10.

In accordance with the present invention, synthetic  
methods are provided comprising the steps of:

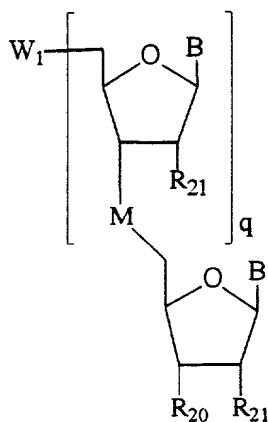
(a) providing a compound of formula IA, IB, IC or ID:



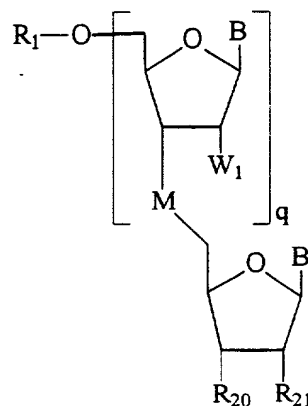
IA



IB



IC



ID

wherein:

$W_1$  is a linking group, O, NH or S, with a linking group being preferred;

5  $R_1$  is H or a hydroxyl protecting group;

B is a nucleobase;

each  $R_{21}$  is H, OH, F, or a group of formula  $Z-R_{22}-(R_{23})_v$ ;

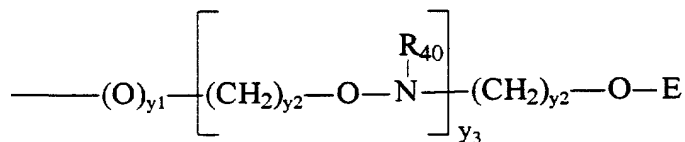
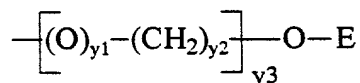
Z is O, S, NH, or  $N-R_{22}-(R_{23})_v$

10  $R_{22}$  is  $C_1$ - $C_{20}$  alkyl,  $C_2$ - $C_{20}$  alkenyl, or  $C_2$ - $C_{20}$  alkynyl;

$R_{23}$  is hydrogen, amino, halogen, hydroxyl, thiol, keto, carboxyl, nitro, nitroso, nitrile, trifluoromethyl, trifluoromethoxy, O-alkyl, S-alkyl, NH-alkyl, N-dialkyl, O-aryl, S-aryl, NH-aryl, O-aralkyl, S-aralkyl, NH-aralkyl,

amino, N-phthalimido, imidazole, azido, hydrazino, hydroxylamino, isocyanato, sulfoxide, sulfone, sulfide, disulfide, silyl, aryl, heterocycle, carbocycle, intercalator, reporter molecule, conjugate, polyamine, polyamide, polyalkylene glycol, polyether, a group that enhances the pharmacodynamic properties of oligonucleotides, or a group that enhances the pharmacokinetic properties of oligonucleotides;

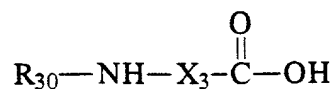
10 v is from 0 to about 10;  
or R<sub>21</sub> has one of the formulas:



wherein:

- y<sub>1</sub> is 0 or 1;  
each y<sub>2</sub> is, independently, 0 to 10;  
15 y<sub>3</sub> is 1 to 10;  
E is N(R<sub>41</sub>)(R<sub>42</sub>) or N=C(R<sub>41</sub>)(R<sub>42</sub>);  
each R<sub>40</sub>, R<sub>41</sub> and R<sub>42</sub> is independently H, C<sub>1</sub>-C<sub>10</sub> alkyl, a nitrogen protecting group, or R<sub>41</sub> and R<sub>42</sub> taken together form a nitrogen protecting group; or R<sub>41</sub> and R<sub>42</sub> taken together with the N or C atom to which they are  
20 attached form a ring structure that can include at least one heterotom selected from N and O;  
q is from 0 to about 50;  
M is an optionally protected internucleoside linkage;
- 25 (b) reacting said compound of formula I with a compound of formula II:



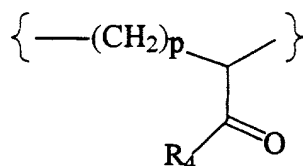


II

wherein:

$R_{30}$  is an amino protecting group;

5  $X_3$  is  $-CH(Z_1)-$  or a group of formula XI:



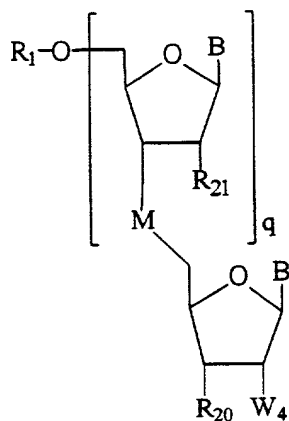
XI

$Z_1$  is the sidechain of a naturally occurring amino acid, or a protected sidechain of a naturally occurring amino acid, wherein the amino acid is preferably glutamic acid;

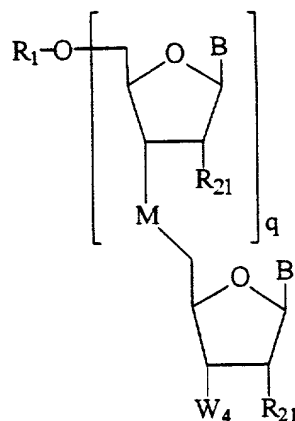
10 acid, wherein the amino acid is preferably glutamic acid;

$p$  is 1 or 2, with 2 being preferred;

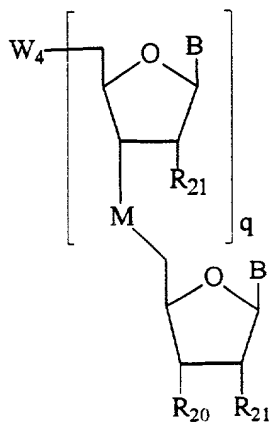
to form a compound of formula IVA, IVB, IVC, or IVD:



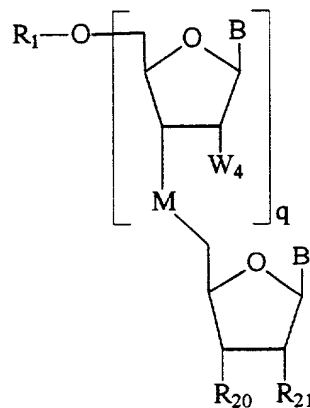
IVA



IVB



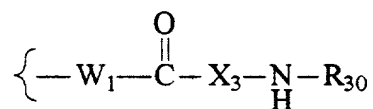
IVC



IVD

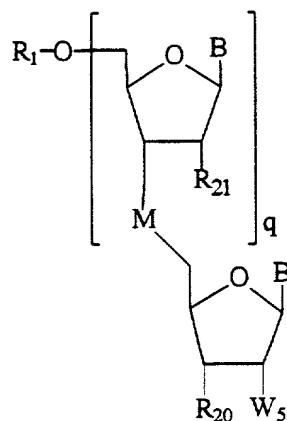
wherein:

$W_4$  has the formula:

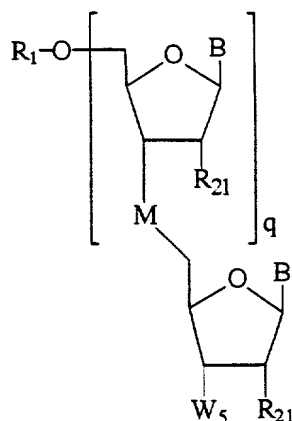


5 and

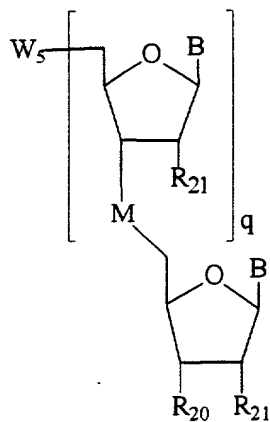
treating said compound of formula IVA, IVB, IVC or IVD with a deprotecting reagent to form a compound of formula VA, VB, VC or VD:



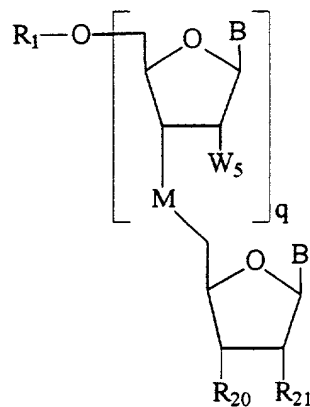
VA



VB

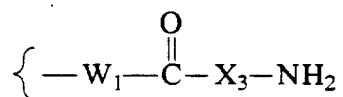


VC

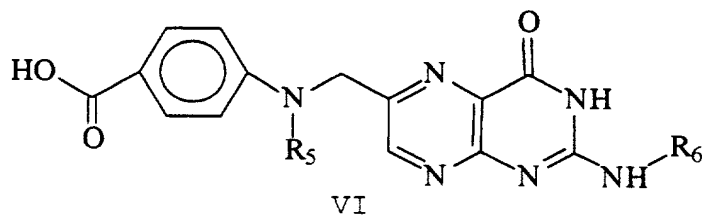


VD

wherein  $W_5$  has the formula:



In some preferred embodiments, the methods of the invention further comprise condensing said compound of formula V with a compound of formula VI:

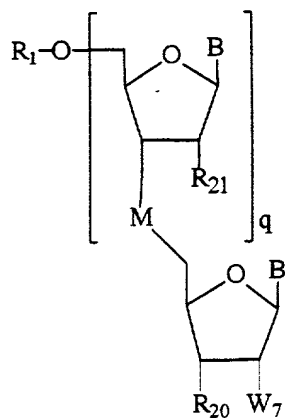


wherein:

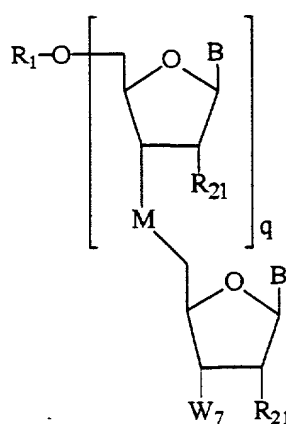
$R_5$  is H or an amino protecting group;

$R_6$  is H or an amino protecting group;

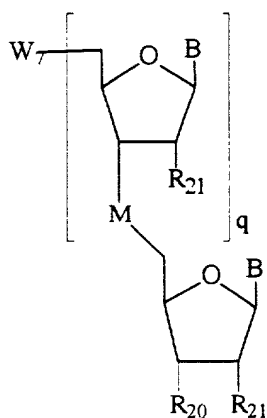
to form a compound of formula VIIA, VIIB, VIIC, OR VIID:



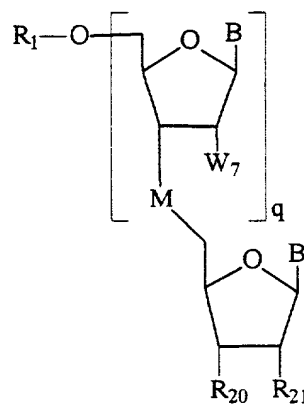
VIIA



VIIB

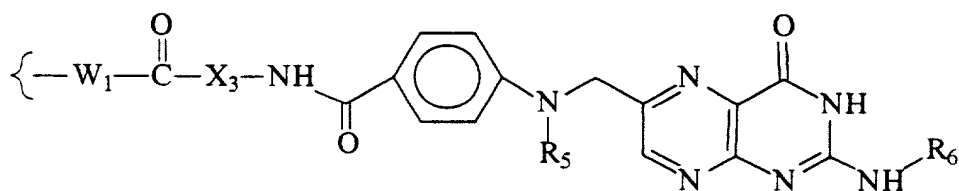


VIIC



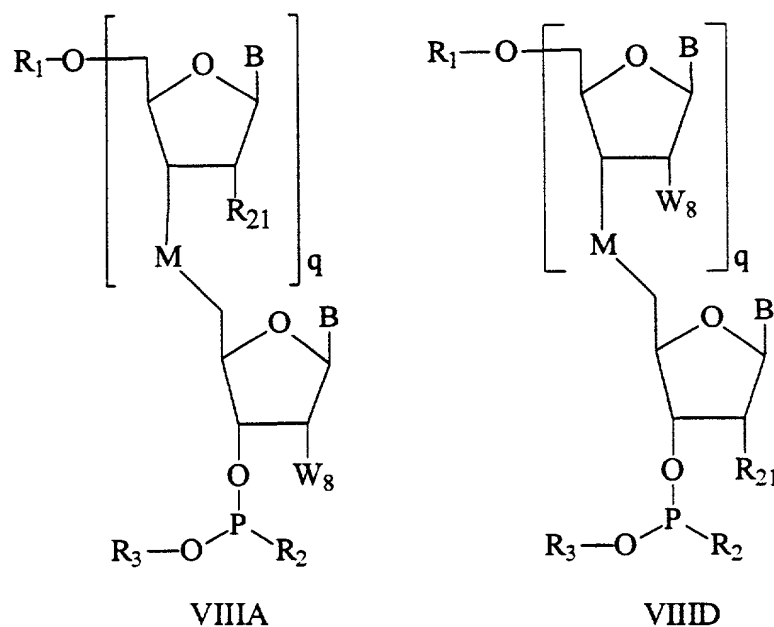
VIID

wherein  $W_7$  has the formula:

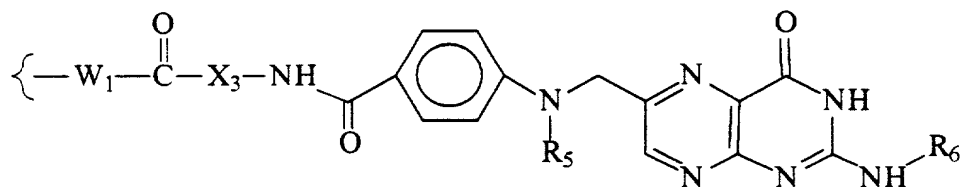


Some preferred embodiments of the methods of the invention further comprise contacting said compound of formula VIIA or VIID with a phosphitylating reagent to form a compound of formula VIIIA or VIIIA-D:

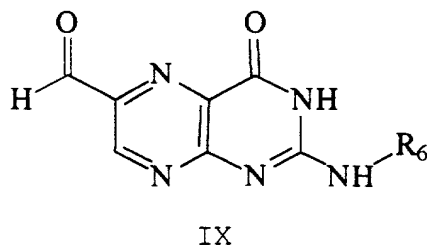
5



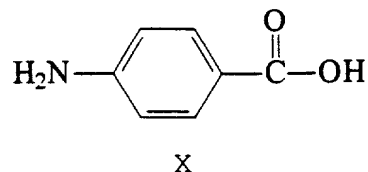
wherein  $W_8$  has the formula:



In some preferred embodiments, said compound of formula VI is prepared by the steps of reacting a compound of  
10 formula IX:

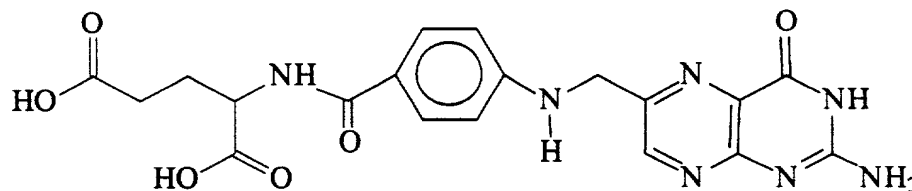


with a compound of formula X:

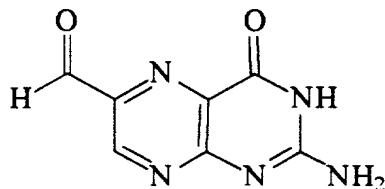


and treating the product of said reaction with a protecting  
5 group reagent to form said compound of formula VI.

Also provided in accordance with the present invention  
are methods for the preparation of a folic acid derivative  
comprising the steps of reacting folic acid:

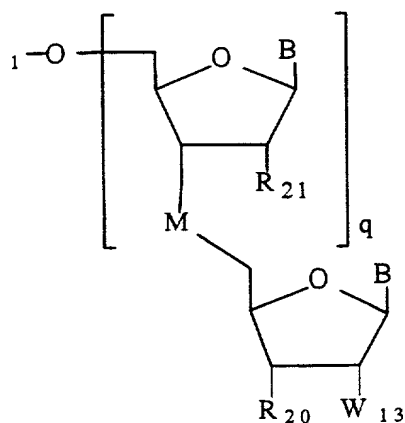


10 with a reagent effective to form pterin aldehyde:

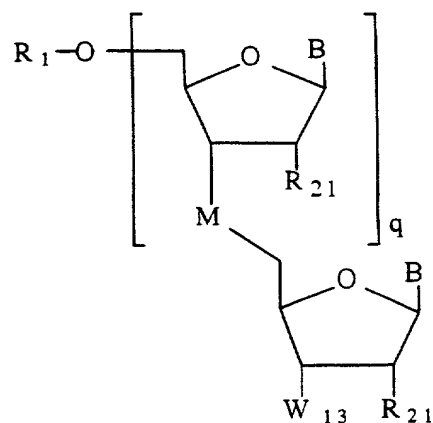


Preferably, the methods further comprise protecting  
the amino group of said pterin aldehyde.

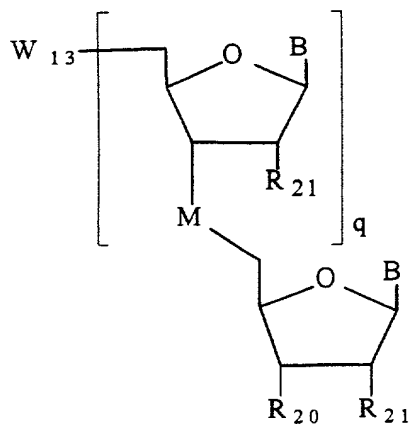
The present invention also provides compounds having  
15 the formula XIII A, XIII B, XIII C or XIII D:



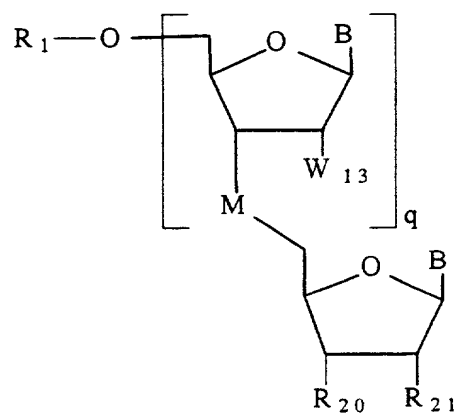
XIII A



XIII B



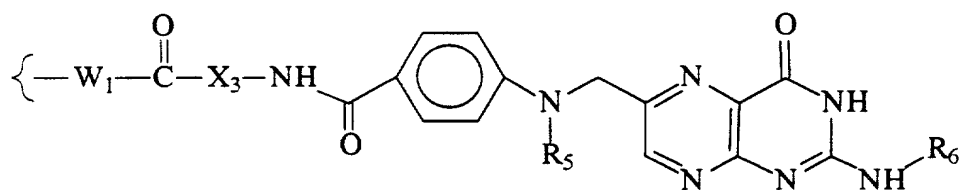
XIII C



XIII D

wherein:

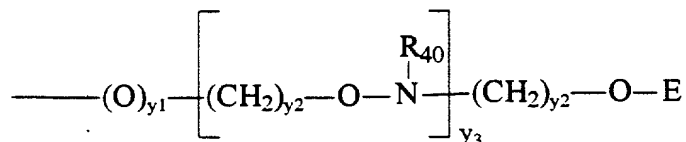
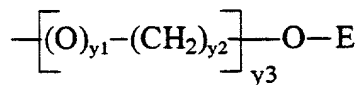
$W_{13}$  has the formula:



- 5  $R_1$  is H or a hydroxyl protecting group;
- $B$  is a nucleobase;
- each  $R_{21}$  is H, OH, F, or a group of formula  $Z-R_{22}-(R_{23})_v$ ;
- $Z$  is O, S, NH or  $N-R_{22}-(R_{23})_v$ ;
- $R_{22}$  is  $C_1-C_{20}$  alkyl,  $C_2-C_{20}$  alkenyl, or  $C_2-C_{20}$
- 10 alkynyl;

R<sub>23</sub> is hydrogen, amino, halogen, hydroxyl, thiol, keto, carboxyl, nitro, nitroso, nitrile, trifluoromethyl, trifluoromethoxy, O-alkyl, S-alkyl, NH-alkyl, N-dialkyl, O-aryl, S-aryl, NH-aryl, O-aralkyl, S-aralkyl, NH-aralkyl, amino, N-phthalimido, imidazole, azido, hydrazino, hydroxylamino, isocyanato, sulfoxide, sulfone, sulfide, disulfide, silyl, aryl, heterocycle, carbocycle, intercalator, reporter molecule, conjugate, polyamine, polyamide, polyalkylene glycol, polyether, a group that enhances the pharmacodynamic properties of oligonucleotides, or a group that enhances the pharmacokinetic properties of oligonucleotides;

or R<sub>21</sub> has one of the formulas:



wherein:

y<sub>1</sub> is 0 or 1;

each y<sub>2</sub> is, independently, 0 to 10;

y<sub>3</sub> is 1 to 10;

E is N(R<sub>41</sub>)(R<sub>42</sub>) or N=C(R<sub>41</sub>)(R<sub>42</sub>);

each R<sub>40</sub>, R<sub>41</sub> and R<sub>42</sub> is independently H, C<sub>1</sub>-C<sub>10</sub> alkyl, a nitrogen protecting group, or R<sub>41</sub> and R<sub>42</sub> taken together form a nitrogen protecting group; or R<sub>41</sub> and R<sub>42</sub> taken together with the N or C atom to which they are attached form a ring structure that can include at least one heterotom selected from N and O;

v is from 0 to about 10;

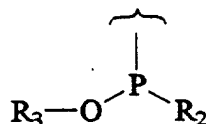
q is from 0 to about 50;

M is an optionally protected internucleoside linkage;



$W_1$  is a linking group, O, NH or S, with a linking group being preferred;

$R_{20}$  is H or a group of formula:



5  $R_2$  is  $-N(R_7)_2$ , or a heterocycloalkyl or heterocycloalkenyl ring containing from 4 to 7 atoms, and having up to 3 heteroatoms selected from nitrogen, sulfur, and oxygen;

$R_7$  is straight or branched chain alkyl having from 10 1 to 10 carbons;

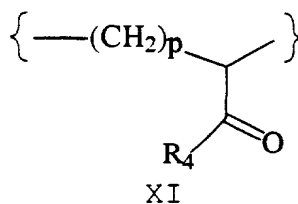
$R_3$  is a phosphorus protecting group;

$n$  is from 1 to about 10;

$R_5$  is H or an amino protecting group;

$R_6$  is H or an amino protecting group;

15  $X_3$  is  $-CH(Z_1)-$  or a group of formula XI:

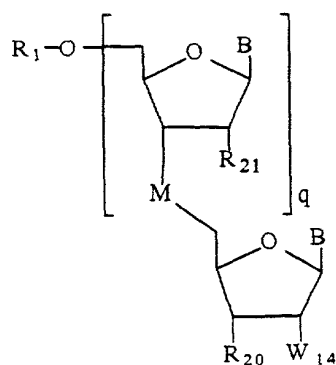


$Z_1$  is the sidechain of a naturally occurring amino acid, or a protected sidechain of a naturally occurring amino acid, wherein the amino acid is preferably glutamic acid;

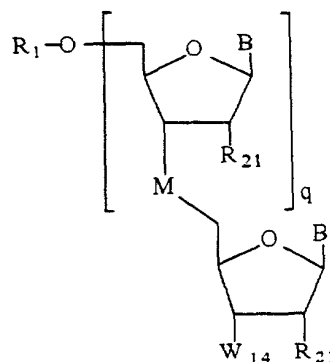
$p$  is 1 or 2, with 2 being preferred; and

$R_4$  is a hydroxyl group, or a protected hydroxyl group.

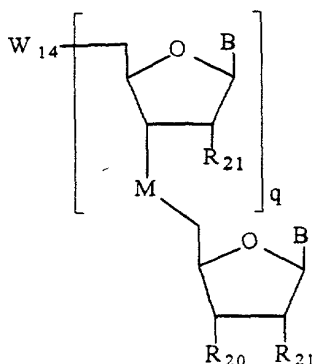
The present invention also provides compositions comprising a compound of formula XIII A-D wherein  $X_3$  has the 25 formula XI and  $p$  is 2, said composition being substantially free of a compound of formula XIVA, XIVB, XIVC, or XIVD:



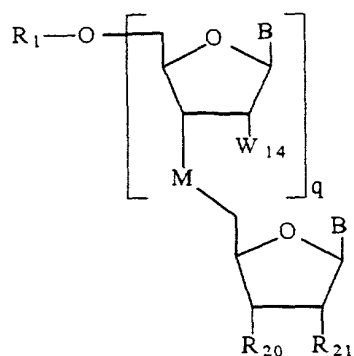
XIVA



XIVB



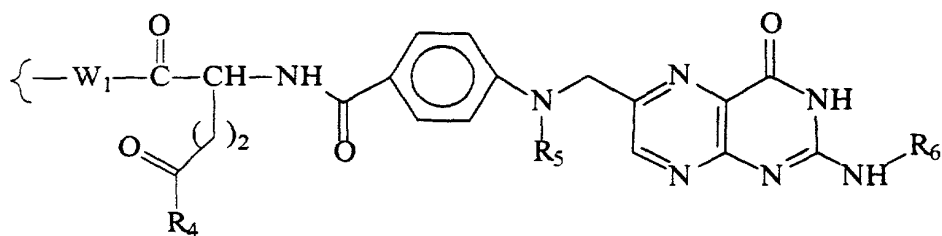
XIVC



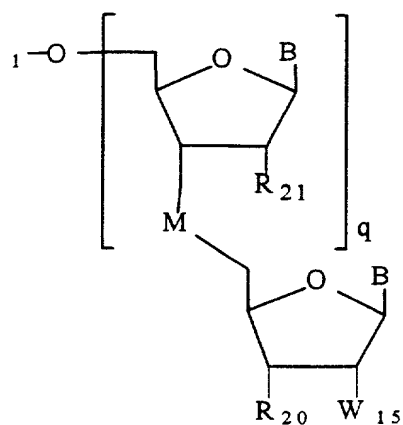
XIVD

wherein:

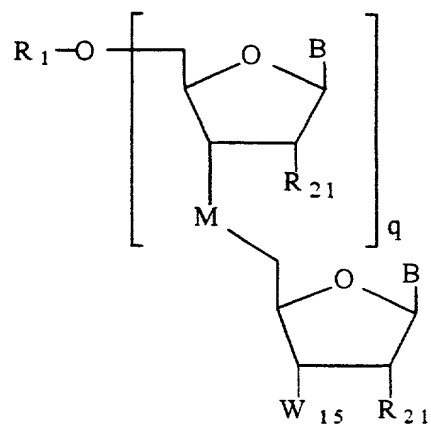
$W_{14}$  has the formula:



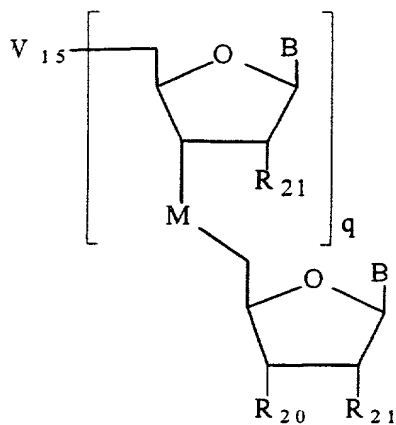
- 5 Also provided in accordance with the present invention are compositions comprising a compound of formula XIIIA-D wherein  $X_3$  has the formula XII and  $m$  is 2, said composition being substantially free of a compound of formula XVA, XVB, XVC or XVD:



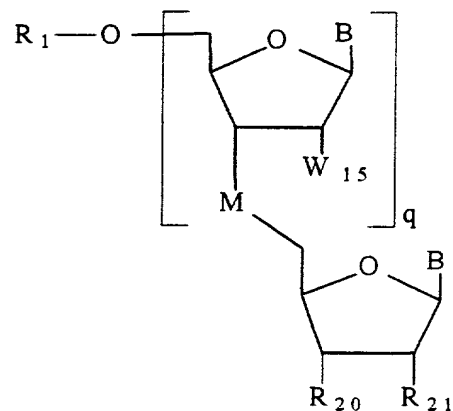
XVA



XVB

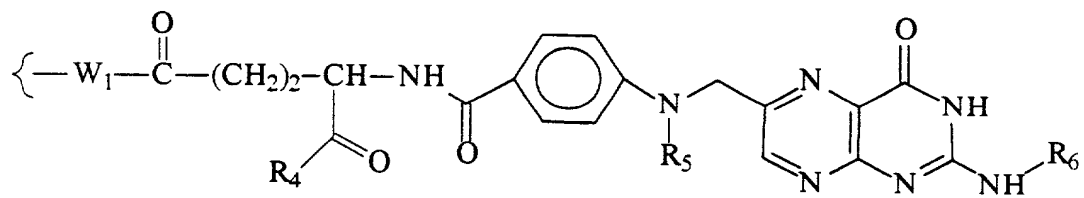


XVC

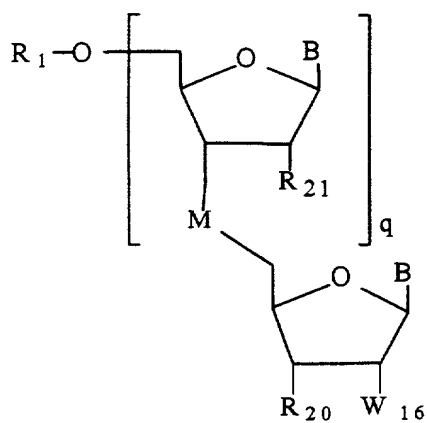


XVD

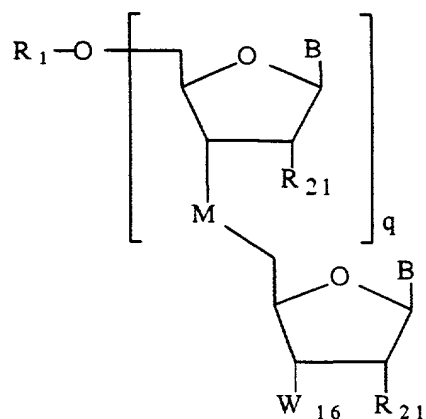
wherein  $W_{15}$  has the formula:



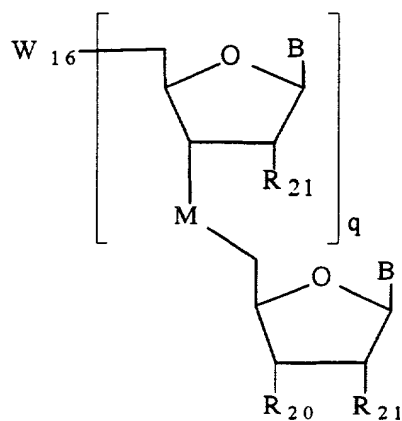
The present invention further provides compounds  
5 having the formula XVIA, XVIB, XVIC or XVID:



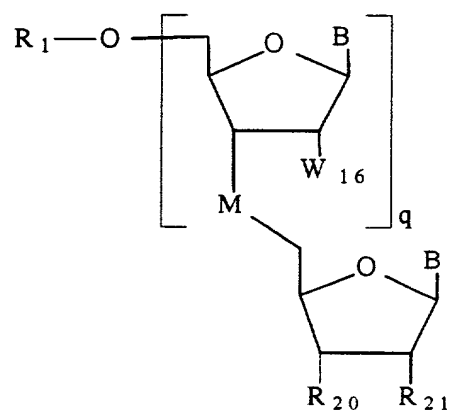
XVI A



XVI B



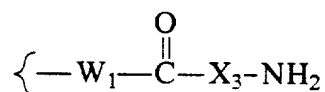
XVI C



XVI D

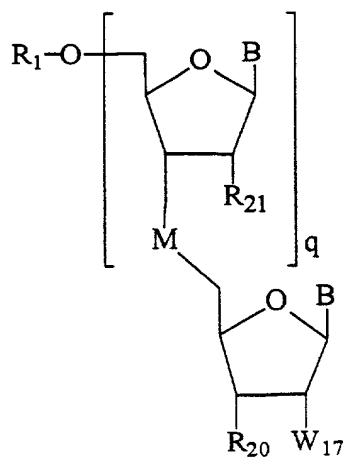
wherein:

$W_{16}$  has the formula:

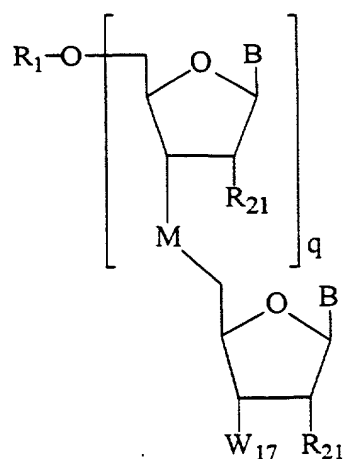


5        wherein the constituent variables are as defined above.

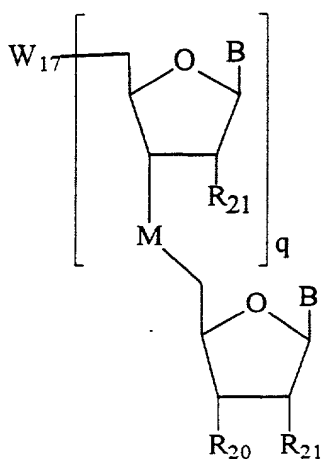
Also provided in accordance with the present invention are compounds having the formula XVI A, XVI B, XVI C or XVI D:



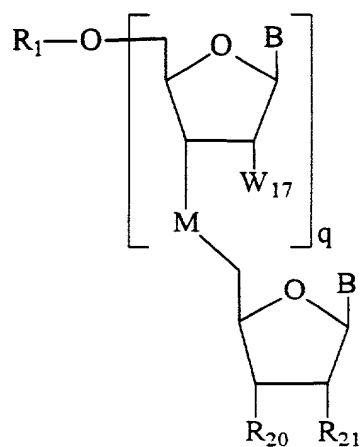
XVIII A



XVIII B



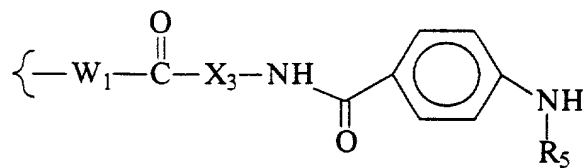
XVIII C



XVIII D

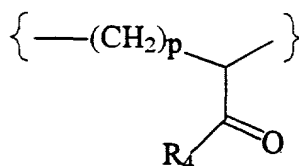
wherein:

$W_{17}$  has the formula:



5 wherein the constituent variables are as defined above.

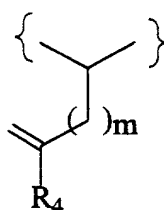
In some preferred embodiments of the foregoing methods and compounds,  $X_3$  is a group of formula XI:



XI

wherein:

- p is 1 or 2, with 2 being preferred;  
 5  $\text{R}_4$  is a hydroxyl group, or a protected hydroxy group;  
 or  $\text{X}_3$  is a group of formula XII:



XII

wherein m is 1 or 2, with 2 being preferred.

- 10 In some preferred embodiments of the foregoing compounds and methods, q is 0.

In further preferred embodiments of the foregoing compounds and methods,  $\text{X}_3$  is a group of formula XI, preferably wherein p is 2.

- 15 In further preferred embodiments of the foregoing compounds and methods,  $\text{X}_3$  is a group of formula XII, preferably wherein m is 2.

- In preferred embodiments of the foregoing compounds and methods,  $\text{W}_1$  has the formula  $-\text{O}-(\text{CH}_2)_n-\text{NH}-$ , where n is from  
 20 1 to about 10, with 6 being preferred.

- In some preferred embodiments of the foregoing compounds and methods,  $\text{R}_1$  is dimethoxytrityl. In further preferred embodiments,  $\text{R}_4$  is t-butoxy. In still further preferred embodiments,  $\text{R}_5$  is trifluoroacetyl. In further  
 25 preferred embodiments,  $\text{R}_6$  is  $-\text{C}(=\text{O})-\text{CH}(\text{CH}_3)_2$ . In yet further preferred embodiments,  $\text{R}_{30}$  is fluorene-9-yl methoxycarbonyl.

In some preferred embodiments of the foregoing compounds and methods,  $\text{R}_1$  is dimethoxytrityl,  $\text{W}_1$  has the

formula  $-O-(CH_2)_n-NH-$  where  $n$  is 6,  $X_3$  has the formula XI where  $p$  is 2,  $R_4$  is *t*-butoxy,  $R_5$  is trifluoroacetyl,  $R_6$  is  $-C(=O)-CH(CH_3)_2$ , and  $R_{30}$  is FMOC.

#### BRIEF DESCRIPTION OF THE DRAWINGS

5        Figure 1 shows the synthesis of  $N^2, N^{10}$ -protected pterioic acid.

Figure 2 shows the synthesis of 2'-O-(g-(a-*t*-butyl-glutamoyl(hexylamino)))-5'-O-DMT-5-methyl-uridine.

10       Figure 3 shows the synthesis of the phosphoramidite of the g-conjugate of folic acid attached to 2'-O-(hexylamino)-5'-O-DMT-5-methyl-uridine.

Figure 4 shows the synthesis of 2'-O-(a-(g-*t*-butyl-glutamoyl(hexylamino)))-5'-O-DMT-5-methyl-uridine.

15       Figure 5 shows the synthesis of the a-conjugate of folic acid attached to 2'-O-(hexylamino)-5'-O-DMT-uridine. shows the synthesis of 2'-O-(g-(a-*t*-butyl-glutamoyl(hexylamino)))-5'-O-DMT-5-methyl-uridine.

20       Figure 6 shows the synthesis of the phosphoramidite of the fully protected g-conjugate of folic acid attached to 2'-O-(hexylamino)-5'-O-DMT-uridine.

Figure 7 shows the synthesis of the phosphoramidite of the fully protected a-conjugate of folic acid attached to 2'-O-(hexylamino)-5'-O-DMT-uridine.

25       Figures 8a and 8b show the synthesis of the g-conjugate of a-O-allyl protected folic acid attached to 2'-O-(hexylamino)-5'-O-DMT-5-methyl-uridine, and the synthesis of the phosphoramidite of the g-conjugate of a-O-allyl protected folic acid attached to 2'-O-(hexylamino)-5'-O-DMT-5-methyl-uridine.

30       Figure 9 is a schematic showing the synthesis of compound 53.

Figure 10 is a schematic showing the synthesis of compounds 57 and 60.

Figure 11 is a schematic showing the synthesis of

compounds 61 and 53.

Figure 12 is a schematic showing the synthesis of compound 64.

Figure 13 is a schematic showing the synthesis of  
5 compound 68.

Figure 14 is a schematic showing the synthesis of compound 69.

Figure 15 is a schematic showing the synthesis of compound 70.

10 Figure 16 is a schematic showing the synthesis of compound 70.

Figure 17 is a schematic showing the synthesis of compound 75.

Figure 18 shows the structure of compound 74b.

15 Figure 19 is a schematic showing the synthesis of compound 78.

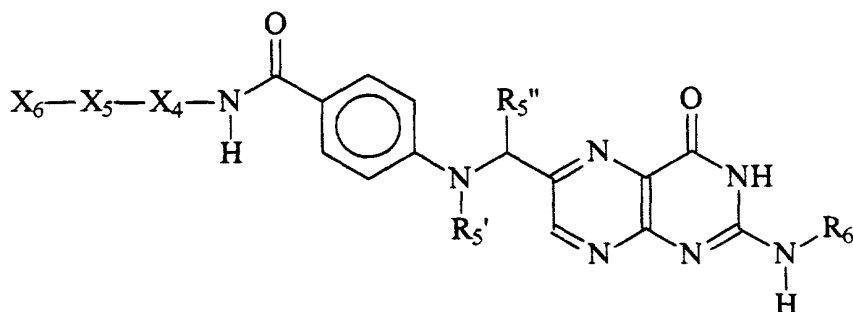
Figure 20 shows representative compounds of the present invention.

#### DETAILED DESCRIPTION OF THE INVENTION

20 The present invention provides nucleosidic and non-nucleosidic folate conjugates. The present invention also provides methods for preparation of nucleosidic and non-nucleosidic folate conjugates as well as oligonucleotide-folate conjugates. The present invention further provides  
25 folate and folate analog-containing conjugates that have improved therapeutic properties, including improved transfer across cellular membranes.

In one preferred embodiment, the present invention provides compounds of formula XVII:

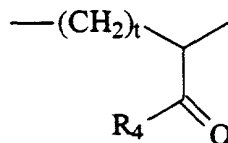




XVII

wherein:

$X_4$  is  $-\text{CH}(X_4)$  or a group of formula:



- 5  $X_4$  is the side chain of a naturally-occurring or non-naturally-occurring amino acid, or a protected side chain of a naturally-occurring or non-naturally-occurring amino acid;
- $X_5$  is  $-\text{N}(X_6)\text{C}(=\text{O})-$ ,  $-\text{C}(=\text{O})\text{NH}-$ ,  $-\text{NHC}(=\text{O})-$ ,  $-\text{OC}(=\text{O})\text{NH}-$ ,  $-\text{C}(=\text{S})\text{NH}-$ ,  $-\text{SC}(=\text{S})\text{NH}-$ ,  $-\text{SC}(=\text{O})\text{NH}-$ ,  $-\text{OC}(=\text{S})\text{NH}-$ ,  $-\text{C}(=\text{O})\text{O}-$ ,  
 10  $-\text{C}(=\text{O})(\text{CH}_2)_n-$  or a bond;
- $n$  is an integer from 1 to 50;
- each  $X_6$  and  $X_6$  is, independently, hydrogen or a hydrocarbyl group selected from  $\text{C}_1$ - $\text{C}_{10}$  alkyl,  $\text{C}_2$ - $\text{C}_{10}$  alkenyl,  $\text{C}_2$ - $\text{C}_{20}$  alkynyl,  $\text{C}_6$ - $\text{C}_{14}$  aryl,  $\text{C}_6$ - $\text{C}_{14}$  aralkyl,  $\text{C}_3$ - $\text{C}_{14}$  cycloalkyl,  
 15  $\text{C}_5$ - $\text{C}_{14}$  fused cycloalkyl,  $\text{C}_4$ - $\text{C}_{14}$  heterocycle,  $\text{C}_4$ - $\text{C}_{14}$  heterocyclalkyl,  $\text{C}_4$ - $\text{C}_{14}$  heteroaryl and  $\text{C}_4$ - $\text{C}_{14}$  heteroarylalkyl; wherein said hydrocarbyl group is substituted with at least two hydroxyl groups, and is optionally substituted with oxo, acyl, alkoxy,  
 20 alkoxy carbonyl, alkyl, alkenyl, alkynyl, amino, amido, azido, aryl, heteroaryl, carboxylic acid, cyano, guanidino, halo, haloalkyl, haloalkoxy, hydrazino, ODMT, alkylsulfonyl, nitro, sulfide, sulfone, sulfonate, sulfonamide, thiol, and thioalkoxy; provided that  $X_6$  is not hydrogen;

$R_4$  is a hydroxyl group or a protected hydroxyl group;

$R_5$  is hydrogen,  $C_1$ - $C_{10}$  alkyl,  $C_2$ - $C_{10}$  alkenyl,  $C_2$ - $C_{20}$  alkynyl,  $C_6$ - $C_{14}$  aryl or an amino protecting group;

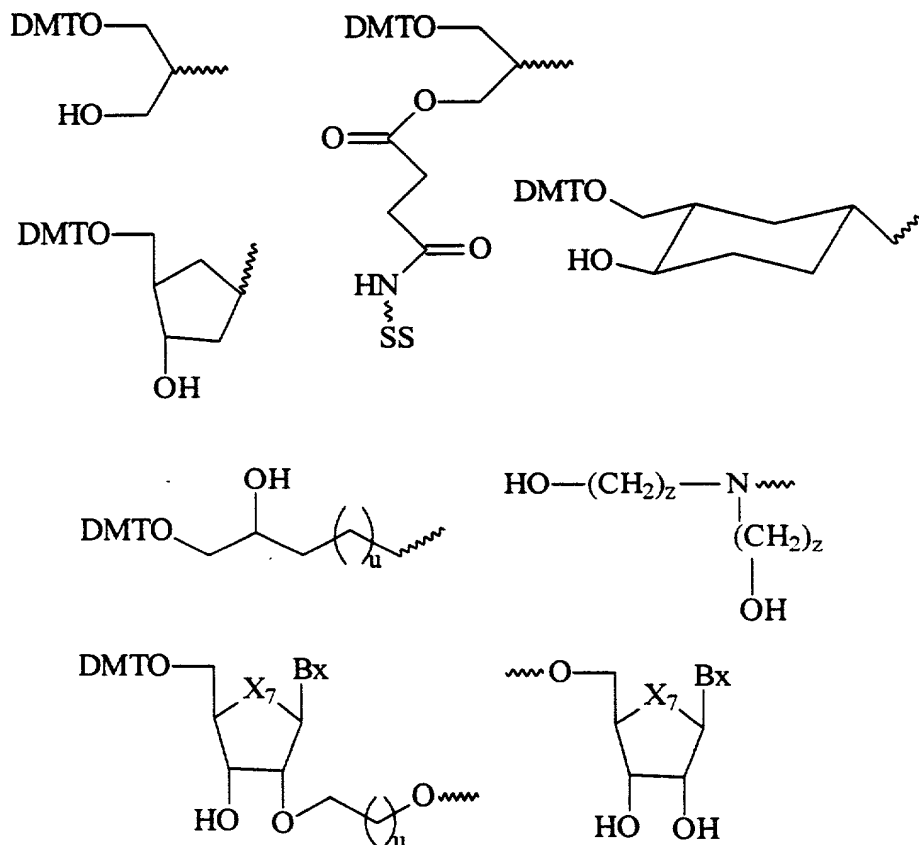
$R_5$  is hydrogen,  $C_1$ - $C_{10}$  alkyl,  $C_2$ - $C_{10}$  alkenyl,  $C_2$ - $C_{20}$  alkynyl,  $C_6$ - $C_{14}$  aryl,  $C_6$ - $C_{14}$  aralkyl,  $C_3$ - $C_{14}$  cycloalkyl,  $C_5$ - $C_{14}$  fused cycloalkyl, formyl, aminoalkyl or hydroxymethyl;

$R_6$  is hydrogen or an amino-protecting group; and

$t$  is 1 or 2.

In a preferred embodiment,  $X_6$  has one of the formulas:

10



wherein:

SS is a solid support;

$X_7$  is O or  $CH_2$ ;

Bx is a nucleobase,  $C_4$ - $C_{14}$  heterocyclyl or hydrogen;

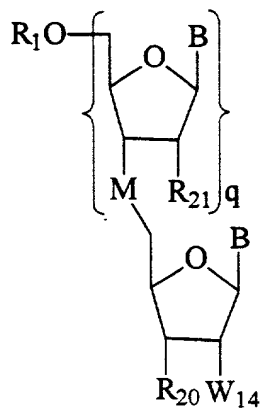
15

$z$  is an integer from 1 to 50; and

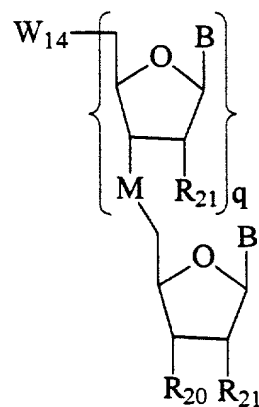
$u$  is an integer from 2 to 5.

The present invention also provides compounds having

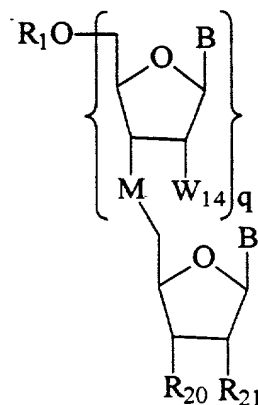
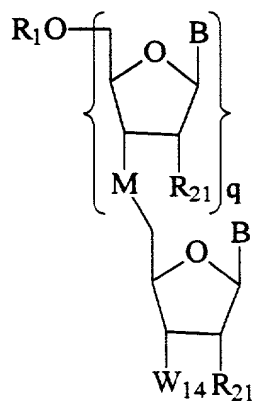
formula XVIA, XVIB, XVIC or XVID:



XVIA

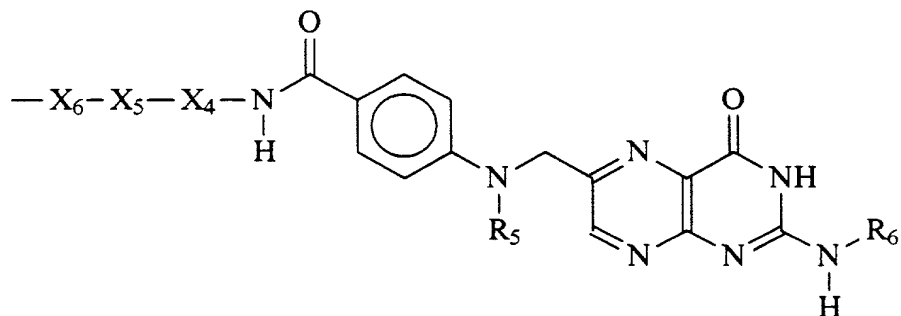


XVIC



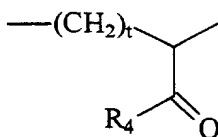
wherein:

$W_{14}$  has the formula:



wherein:

5  $X_4$  is  $-CH(X_4')$  or a group of formula:



$\text{X}_4$  is the side chain of a naturally-occurring or non-naturally-occurring amino acid, or a protected side chain of a naturally-occurring or non-naturally-occurring amino acid;

$t$  is 1 or 2;

- 5  $\text{X}_5$  is  $\text{---N}(\text{X}_6)\text{C}(\text{O})\text{---}$ ,  $\text{---C}(\text{O})\text{NH}\text{---}$ ,  $\text{---NH}(\text{CO})\text{---}$ ,  $\text{---OC}(\text{O})\text{NH}\text{---}$ ,  $\text{---C}(\text{S})\text{NH}\text{---}$ ,  $\text{---SC}(\text{S})\text{NH}\text{---}$ ,  $\text{---SC}(\text{O})\text{NH}\text{---}$ ,  $\text{---OC}(\text{S})\text{NH}\text{---}$ ,  $\text{---C}(\text{O})\text{O}\text{---}$ ,  $\text{---C}(\text{O})(\text{CH}_2)_n\text{---}$  or a bond;

$n$  is an integer from 1 to 50;

- each  $\text{X}_5$  and  $\text{X}_6$  is, independently, a bond, hydrogen or  
 10 a hydrocarbyl group selected from  $\text{C}_1\text{--C}_{10}$  alkyl,  $\text{C}_2\text{--C}_{10}$  alkenyl,  $\text{C}_2\text{--C}_{20}$  alkynyl,  $\text{C}_6\text{--C}_{14}$  aryl,  $\text{C}_6\text{--C}_{14}$  aralkyl,  $\text{C}_3\text{--C}_{14}$  cycloalkyl,  $\text{C}_5\text{--C}_{14}$  fused cycloalkyl,  $\text{C}_4\text{--C}_{14}$  heterocycle,  $\text{C}_4\text{--C}_{14}$  heterocyclylalkyl,  $\text{C}_4\text{--C}_{14}$  heteroaryl and  $\text{C}_4\text{--C}_{14}$  heteroarylalkyl; wherein said hydrocarbyl group is  
 15 substituted with at least two hydroxyl groups, and is optionally substituted with oxo, acyl, alkoxy, alkoxy carbonyl, alkyl, alkenyl, alkynyl, amino, amido, azido, aryl, heteroaryl, carboxylic acid, cyano, guanidino, halo, haloalkyl, haloalkoxy, hydrazino, ODMT, alkylsulfonyl, nitro,  
 20 sulfide, sulfone, sulfonate, sulfonamide, thiol, and thioalkoxy; provided that  $\text{X}_5$  is not hydrogen and  $\text{X}_6$  is not a bond;

$\text{R}_1$  is hydrogen or a hydroxyl protecting group;

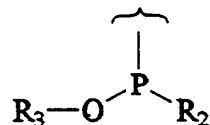
$\text{R}_4$  is a hydroxyl group or a protected hydroxyl group;

- 25 each  $\text{R}_5$  and  $\text{R}_{40}$  is, independently, hydrogen,  $\text{C}_1\text{--C}_{10}$  alkyl,  $\text{C}_2\text{--C}_{10}$  alkenyl,  $\text{C}_2\text{--C}_{20}$  alkynyl,  $\text{C}_6\text{--C}_{14}$  aryl or an amino-protecting group;

- $\text{R}_{50}$  is hydrogen,  $\text{C}_1\text{--C}_{10}$  alkyl,  $\text{C}_2\text{--C}_{10}$  alkenyl,  $\text{C}_2\text{--C}_{20}$  alkynyl,  $\text{C}_6\text{--C}_{14}$  aryl,  $\text{C}_6\text{--C}_{14}$  aralkyl,  $\text{C}_3\text{--C}_{14}$  cycloalkyl,  $\text{C}_5\text{--C}_{14}$  fused cycloalkyl, formyl, aminoalkyl or hydroxymethyl;  
 30

$R_6$  is hydrogen or an amino protecting group;

$R_{20}$  is hydrogen or a group of formula:



$R_2$  is  $-N(R_7)_2$ , or a heterocycloalkyl or  
5 heterocycloalkenyl ring containing from 4 to 7 atoms, and  
having up to 3 heteroatoms selected from nitrogen, sulfur,  
and oxygen;

$R_7$  is straight or branched chain alkyl having from 1 to  
10 carbons;

10  $R_3$  is a phosphorus protecting group;

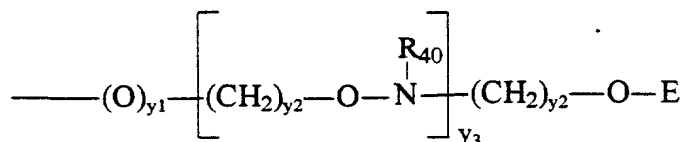
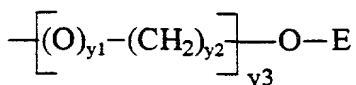
$R_{21}$  is hydrogen, hydroxyl, fluoro or a group of formula  
 $Z-R_{22}-(R_{23})_v$ ;

$Z$  is O, S, NH or  $N-R_{22}-(R_{23})_v$ ;

$R_{22}$  is  $C_1-C_{20}$  alkyl,  $C_2-C_{20}$  alkenyl, or  $C_2-C_{20}$  alkynyl;

15  $R_{23}$  is hydrogen, amino, halogen, hydroxyl, thiol, keto,  
carboxyl, nitro, nitroso, nitrile, trifluoromethyl,  
trifluoromethoxy, O-alkyl, S-alkyl, NH-alkyl, N-dialkyl, O-  
aryl, S-aryl, NH-aryl, O-aralkyl, S-aralkyl, NH-aralkyl,  
amino, N-phthalimido, imidazole, azido, hydrazino,  
20 hydroxylamino, isocyanato, sulfoxide, sulfone, sulfide,  
disulfide, silyl, aryl, heterocycle, carbocycle, inter-  
calator, reporter molecule, conjugate, polyamine, polyamide,  
polyalkylene glycol, polyether, a group that enhances the  
pharmacodynamic properties of oligonucleotides, or a group  
25 that enhances the pharmacokinetic properties of oligonucleo-  
tides;

or  $R_{21}$  has one of the formulas:



wherein:

y<sub>1</sub> is 0 or 1;

each y<sub>2</sub> is, independently, 0 to 10;

5 y<sub>3</sub> is 1 to 10;

E is N(R<sub>41</sub>)(R<sub>42</sub>) or N=C(R<sub>41</sub>)(R<sub>42</sub>);

each R<sub>41</sub> and each R<sub>42</sub> is independently H, C<sub>1</sub>-C<sub>10</sub> alkyl, a nitrogen protecting group, or R<sub>41</sub> and R<sub>42</sub> taken together form a nitrogen protecting group; or R<sub>41</sub> and R<sub>42</sub> taken together with  
 10 the N or C atom to which they are attached form a ring structure that can include at least one heterotom selected from N and O;

B is a nucleobase;

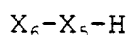
M is an optionally protected internucleoside linkage;

15 q is from zero to about 50; and

v is from zero to about 10.

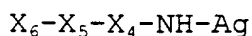
The present invention also provides methods for the preparation of compounds of formula XVII comprising the steps of:

20 (a) providing a hydroxy compound of formula:



(b) protecting said hydroxyl groups of X<sub>6</sub> with protecting groups to form a protected hydroxy compound;

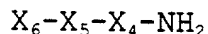
(c) reacting said protected hydroxy compound with an  
 25 amino-protected amino acid to form a covalently linked hydroxy compound of formula:



wherein Ag is an amino protecting group;

(d) cleaving the amino-protecting group of said  
 30 covalently linked hydroxy compound to form a hydroxy compound

bearing a deprotected amino group and having formula:



(e) reacting said amino group with a folate moiety;  
and

- 5 (f) cleaving the protecting groups on said hydroxyl groups of step (b) to form a compound of formula XVII.

The methods of the present invention also provide compounds of the formula XVII wherein  $R_4$  is a hydroxyl group protected with  $C_1$ - $C_{10}$  alkyl,  $C_2$ - $C_{10}$  alkenyl or  $C_2$ - $C_{20}$  alkynyl,  
10 the methods further comprising the step of deprotecting the  $R_4$  group with a deprotecting reagent. In a preferred embodiment,  $R_4$  is a methoxy group or an allyl protected hydroxyl group.

This deprotection step is performed by the use of  
15 aqueous amines. Preferred amines include, but are not limited to, piperidine, piperazine, pyrrolidine and morpholine. Deprotection may also be achieved using a combination of a mercapto compound and an amine.

In a preferred embodiment of the present invention the  
20 mercapto compound is mercaptoethanol, thiocresol, benzyl mercaptan, thiophenol, cysteine, cysteamine, glutathione or dithiothreitol.

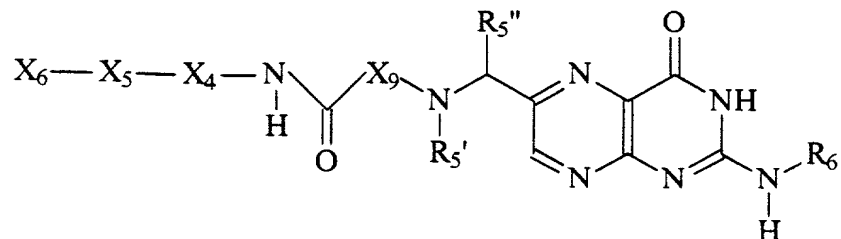
The deprotection reagents lead to an unexpected increase in the overall synthetic yield. This increase in  
25 yield is due, in part, to the use of a mercapto nucleophile (mercapto compound) in combination with ammonia in the deblocking step. Although not wanting to be bound by theory, it is thought that nucleophilic attack, by the more nucleophilic mercapto compound enhances the overall  
30 efficiency of the deblocking step. This enhanced reaction efficiency allows performance of the deblocking step at a lower temperature than has been previously reported.

The present invention also provides methods further comprising the steps of:

(g) protecting one of said hydroxyl groups of  $X_6$  with a dimethoxytrityl group; and

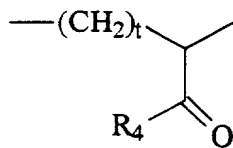
(h) phosphitylating the other of said hydroxyl groups of  $X_6$ .

5 The present invention also provides compounds of the formula:



wherein:

$X_4$  is  $-\text{CH}(X_4)-$  or a group of formula:



$X_4$  is the side chain of a naturally-occurring or non-naturally-occurring amino acid, or a protected side chain of a naturally-occurring or non-naturally-occurring amino acid;

$X_5$  is  $-\text{N}(X_5)\text{C}(=\text{O})-$ ,  $-\text{C}(=\text{O})\text{NH}-$ ,  $-\text{NHC}(=\text{O})-$ ,  $-\text{OC}(=\text{O})\text{NH}-$ ,  $-\text{C}(=\text{S})\text{NH}-$ ,  $-\text{SC}(=\text{S})\text{NH}-$ ,  $-\text{SC}(=\text{O})\text{NH}-$ ,  $-\text{OC}(=\text{S})\text{NH}-$ ,  $-\text{C}(=\text{O})\text{O}-$ ,  $-\text{C}(=\text{O})(\text{CH}_2)_n-$  or a bond;

15  $n$  is an integer from 1 to 50;

each  $X_6$ ,  $X_6$ , and  $X_9$  is, independently, a bond, hydrogen or a hydrocarbyl group selected from  $\text{C}_1$ - $\text{C}_{10}$  alkyl,  $\text{C}_2$ - $\text{C}_{10}$  alkenyl,  $\text{C}_2$ - $\text{C}_{20}$  alkynyl,  $\text{C}_6$ - $\text{C}_{14}$  aryl,  $\text{C}_6$ - $\text{C}_{14}$  aralkyl,  $\text{C}_3$ - $\text{C}_{14}$  cycloalkyl,  $\text{C}_5$ - $\text{C}_{14}$  fused cycloalkyl,  $\text{C}_4$ - $\text{C}_{14}$  heterocycle,  $\text{C}_4$ - $\text{C}_{14}$  heterocyclalkyl,  $\text{C}_4$ - $\text{C}_{14}$  heteroaryl and  $\text{C}_4$ - $\text{C}_{14}$  heteroarylalkyl; wherein said hydrocarbyl group is substituted with at least two hydroxyl groups, and is optionally substituted with oxo, acyl, alkoxy, alkoxy carbonyl, alkyl, alkenyl, alkynyl, amino, amido, azido, 25 aryl, heteroaryl, carboxylic acid, cyano, guanidino, halo,



haloalkyl, haloalkoxy, hydrazino, hydroxyl, ODMT, alkylsulfonyl, nitro, sulfide, disulfide, sulfone, sulfonate, sulfonamide, thiol, and thioalkoxy; provided that each of  $X_5$  and  $X_6$  is not a bond, and  $X_6$  is not hydrogen;

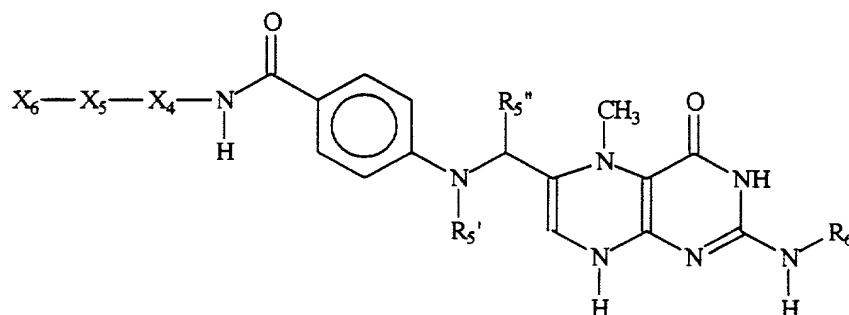
5  $R_4$  is a hydroxyl group or a protected hydroxyl group;

$R_5$  is hydrogen,  $C_1$ - $C_{10}$  alkyl,  $C_2$ - $C_{10}$  alkenyl,  $C_2$ - $C_{20}$  alkynyl,  $C_6$ - $C_{14}$  aryl or an amino-protecting group

$R_{5'}$  is hydrogen,  $C_1$ - $C_{10}$  alkyl,  $C_2$ - $C_{10}$  alkenyl,  $C_2$ - $C_{20}$  alkynyl,  $C_6$ - $C_{14}$  aryl,  $C_6$ - $C_{14}$  aralkyl,  $C_3$ - $C_{14}$  cycloalkyl, formyl, 10 aminoalkyl or hydroxymethyl;

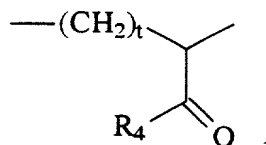
$R_6$  is hydrogen or an amino-protecting group; and  $t$  is 1 or 2.

The present invention further provides compounds of the formula:



15 wherein:

$X_4$  is  $-CH(X_4')$  or a group of formula:



$X_4'$  is the side chain of a naturally-occurring or non-naturally-occurring amino acid, or a protected side chain of a naturally-occurring or non-naturally-occurring amino acid;

20  $X_5$  is  $-N(X_6)C(O)-$ ,  $-C(O)NH-$ ,  $-NHC(O)-$ ,  $-OC(O)NH-$ ,  $-C(S)NH-$ ,  $-SC(S)NH-$ ,  $-SC(O)NH-$ ,  $-OC(S)NH-$ ,  $-C(O)O-$ ,  $-C(O)(CH_2)_n-$  or a bond;

$n$  is an integer from 1 to 50;

each of  $X_6$  and  $X_6$  is, independently, hydrogen or a hydrocarbyl group selected from  $C_1$ - $C_{10}$  alkyl,  $C_2$ - $C_{10}$  alkenyl,  $C_2$ - $C_{20}$  alkynyl,  $C_6$ - $C_{14}$  aryl,  $C_6$ - $C_{14}$  aralkyl,  $C_3$ - $C_{14}$  cycloalkyl,  $C_5$ - $C_{14}$  fused cycloalkyl,  $C_4$ - $C_{14}$  heterocycle,  $C_4$ - $C_{14}$  heterocyclalkyl,  $C_4$ - $C_{14}$  heteroaryl and  $C_4$ - $C_{14}$  heteroarylalkyl; wherein said hydrocarbyl group is substituted with at least two hydroxyl groups, and is optionally substituted with oxo, acyl, alkoxy, alkoxy carbonyl, alkyl, alkenyl, alkynyl, amino, amido, azido, aryl, heteroaryl, carboxylic acid, cyano, guanidino, halo, haloalkyl, haloalkoxy, hydrazino, ODMT, alkylsulfonyl, nitro, sulfide, disulfide, sulfone, sulfonate, sulfonamide, thiol, and thioalkoxy; provided that  $X_6$  is not hydrogen;

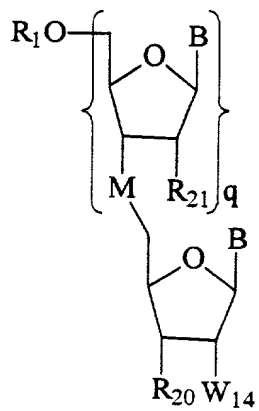
$R_4$  is a hydroxyl group or a protected hydroxyl group;  
 $R_5$  is hydrogen,  $C_1$ - $C_{10}$  alkyl,  $C_2$ - $C_{10}$  alkenyl,  $C_2$ - $C_{20}$  alkynyl,  $C_6$ - $C_{14}$  aryl or an amino-protecting group

$R_5$  is hydrogen,  $C_1$ - $C_{10}$  alkyl,  $C_2$ - $C_{10}$  alkenyl,  $C_2$ - $C_{20}$  alkynyl,  $C_6$ - $C_{14}$  aryl,  $C_6$ - $C_{14}$  aralkyl,  $C_3$ - $C_{14}$  cycloalkyl, formyl, aminoalkyl or hydroxymethyl;

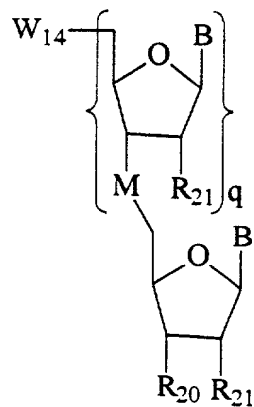
$R_6$  is hydrogen or an amino protecting group; and  $t$  is 1 or 2.

In another preferred embodiment, the present invention provides methods for the preparation of mononucleosides or oligonucleotides having one or more folate molecules conjugated thereto, optionally via one or more linking groups.

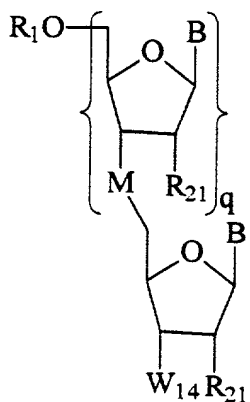
The present invention also provides compounds of the formula XVIA, XVIB, XVIC or XVID:



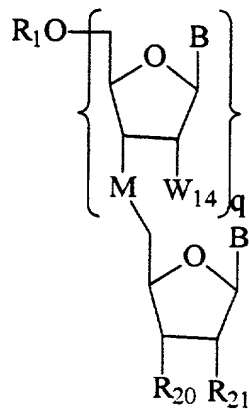
XVIA



XVICS



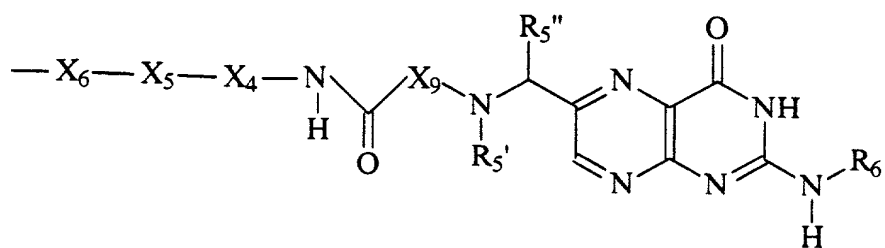
XVIB



XVID

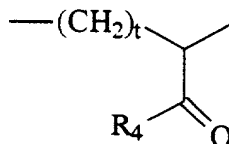
wherein:

$W_{14}$  has the formula:



wherein:

5           X<sub>4</sub> is -CH(X<sub>4</sub>,) or a group of formula:



$\text{X}_4$  is the side chain of a naturally-occurring or non-naturally-occurring amino acid, or a protected side chain of a naturally-occurring or non-naturally-occurring amino acid;

5  $t$  is 1 or 2;

$\text{X}_5$  is  $-\text{N}(\text{X}_6)\text{C}(\text{O})-$ ,  $-\text{C}(\text{O})\text{NH}-$ ,  $-\text{NHC}(\text{O})-$ ,  $-\text{OC}(\text{O})\text{NH}-$ ,  $-\text{C}(\text{S})\text{NH}-$ ,  $-\text{SC}(\text{S})\text{NH}-$ ,  $-\text{SC}(\text{O})\text{NH}-$ ,  $-\text{OC}(\text{S})\text{NH}-$ ,  $-\text{C}(\text{O})\text{O}-$ ,  $-\text{C}(\text{O})(\text{CH}_2)_n-$  or a bond;

$n$  is an integer from 1 to 50;

10 each  $\text{X}_6$ ,  $\text{X}_6$ , and  $\text{X}_9$  is, independently, a bond, hydrogen or a hydrocarbyl group selected from  $\text{C}_1$ - $\text{C}_{10}$  alkyl,  $\text{C}_2$ - $\text{C}_{10}$  alkenyl,  $\text{C}_2$ - $\text{C}_{20}$  alkynyl,  $\text{C}_6$ - $\text{C}_{14}$  aryl,  $\text{C}_6$ - $\text{C}_{14}$  aralkyl,  $\text{C}_3$ - $\text{C}_{14}$  cycloalkyl,  $\text{C}_5$ - $\text{C}_{14}$  fused cycloalkyl,  $\text{C}_4$ - $\text{C}_{14}$  heterocycle,  $\text{C}_4$ - $\text{C}_{14}$  heterocyclylalkyl,  $\text{C}_4$ - $\text{C}_{14}$  heteroaryl and  $\text{C}_4$ - $\text{C}_{14}$  heteroarylalkyl; wherein said hydrocarbyl group is substituted with at least two hydroxyl groups, and is optionally substituted with oxo, acyl, alkoxy, alkoxy carbonyl, alkyl, alkenyl, alkynyl, amino, amido, azido, aryl, heteroaryl, carboxylic acid, cyano, guanidino, halo, haloalkyl, haloalkoxy, hydrazino, ODMT, alkylsulfonyl, nitro, sulfide, disulfide, sulfone, sulfonate, sulfonamide, thiol, and thioalkoxy; provided that each  $\text{X}_6$  and  $\text{X}_9$  is not hydrogen and  $\text{X}_6$  is not a bond;

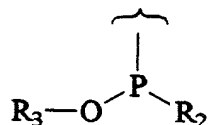
$\text{R}_1$  is hydrogen or a hydroxyl protecting group;

25  $\text{R}_4$  is a hydroxyl group or a protected hydroxyl group; each  $\text{R}_5$  and  $\text{R}_{40}$  is, independently, hydrogen,  $\text{C}_1$ - $\text{C}_{10}$  alkyl,  $\text{C}_2$ - $\text{C}_{10}$  alkenyl,  $\text{C}_2$ - $\text{C}_{20}$  alkynyl,  $\text{C}_6$ - $\text{C}_{14}$  aryl or an amino-protecting group

30  $\text{R}_{50}$  is hydrogen,  $\text{C}_1$ - $\text{C}_{10}$  alkyl,  $\text{C}_2$ - $\text{C}_{10}$  alkenyl,  $\text{C}_2$ - $\text{C}_{20}$  alkynyl,  $\text{C}_6$ - $\text{C}_{14}$  aryl,  $\text{C}_6$ - $\text{C}_{14}$  aralkyl,  $\text{C}_3$ - $\text{C}_{14}$  cycloalkyl, formyl, aminoalkyl or hydroxymethyl;

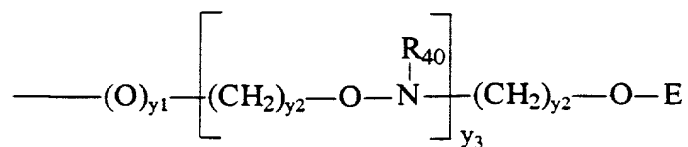
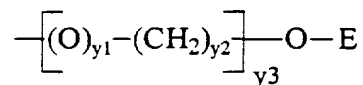
$R_6$  is hydrogen or an amino protecting group;

$R_{20}$  is hydrogen or a group of formula:



- $R_2$  is  $-N(R_7)_2$ , or a heterocycloalkyl or
- 5 heterocycloalkenyl ring containing from 4 to 7 atoms, and having up to 3 heteroatoms selected from nitrogen, sulfur, and oxygen;
- $R_7$  is straight or branched chain alkyl having from 1 to 10 carbons;
- 10  $R_3$  is a phosphorus protecting group;
- $R_{21}$  is hydrogen, hydroxyl, fluoro or a group of formula  $Z-R_{22}-(R_{23})_v$ ;
- $Z$  is O, S, NH or  $N-R_{22}-(R_{23})_v$ ;
- $R_{22}$  is  $C_1-C_{20}$  alkyl,  $C_2-C_{20}$  alkenyl, or  $C_2-C_{20}$  alkynyl;
- 15  $R_{23}$  is hydrogen, amino, halogen, hydroxyl, thiol, keto, carboxyl, nitro, nitroso, nitrile, trifluoromethyl, trifluoromethoxy, O-alkyl, S-alkyl, NH-alkyl, N-dialkyl, O-aryl, S-aryl, NH-aryl, O-aralkyl, S-aralkyl, NH-aralkyl, amino, N-phthalimido, imidazole, azido, hydrazino,
- 20 hydroxylamino, isocyanato, sulfoxide, sulfone, sulfide, disulfide, silyl, aryl, heterocycle, carbocycle, intercalator, reporter molecule, conjugate, polyamine, polyamide, polyalkylene glycol, polyether, a group that enhances the pharmacodynamic properties of oligonucleotides, or a group
- 25 that enhances the pharmacokinetic properties of oligonucleotides;

or  $R_{21}$  has one of the formulas:



wherein:

y<sub>1</sub> is 0 or 1;

each y<sub>2</sub> is, independently, 0 to 10;

5 y<sub>3</sub> is 1 to 10;

E is N(R<sub>41</sub>)(R<sub>42</sub>) or N=C(R<sub>41</sub>)(R<sub>42</sub>);

each R<sub>41</sub> and each R<sub>42</sub> is independently H, C<sub>1</sub>-C<sub>10</sub> alkyl, a  
nitrogen protecting group, or R<sub>41</sub> and R<sub>42</sub> taken together form a  
nitrogen protecting group; or R<sub>41</sub> and R<sub>42</sub> taken together with  
10 the N or C atom to which they are attached form a ring  
structure that can include at least one heterotom selected  
from N and O;

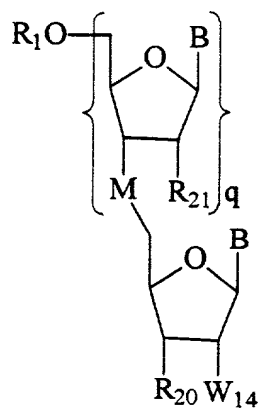
B is a nucleobase;

M is an optionally protected internucleoside linkage;

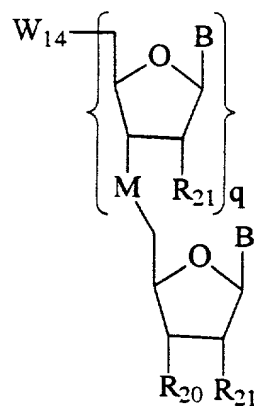
15 q is from zero to about 50; and

v is from zero to about 10.

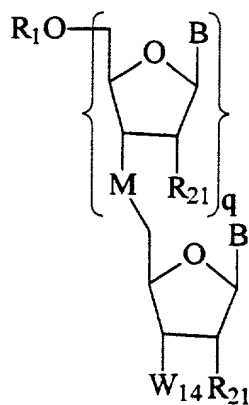
The present invention further provides compounds of  
the formula XVIA, XVIB, XVIC or XVID:



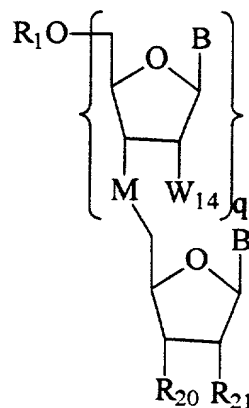
XVI A



XVI C



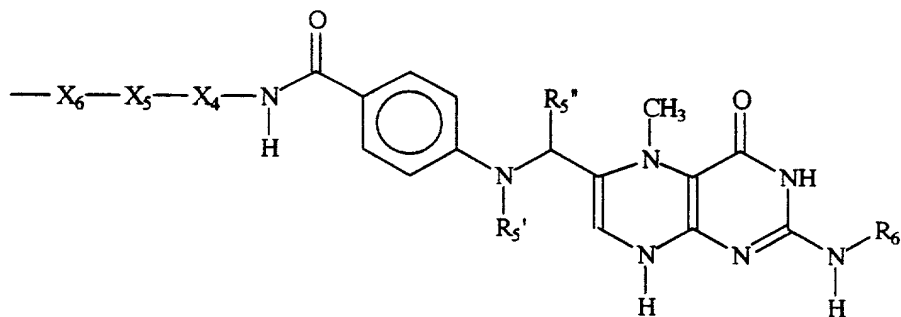
XVI B



XVI D

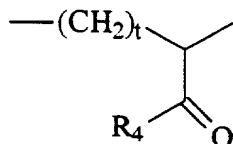
wherein:

$W_{14}$  has the formula:



wherein:

$X_4$  is  $-\text{CH}(X_4)$  or a group of formula:



$X_4$  is the side chain of a naturally-occurring or non-naturally-occurring amino acid, or a protected side chain of a naturally-occurring or non-naturally-occurring amino acid;

$t$  is 1 or 2;

$X_5$  is  $-\text{N}(X_6)\text{C}(\text{O})-$ ,  $-\text{C}(\text{O})\text{NH}-$ ,  $-\text{NHC}(\text{O})-$ ,  $-\text{OC}(\text{O})\text{NH}-$ ,  $-\text{C}(\text{S})\text{NH}-$ ,  $-\text{SC}(\text{S})\text{NH}-$ ,  $-\text{SC}(\text{O})\text{NH}-$ ,  $-\text{OC}(\text{S})\text{NH}-$ ,  $-\text{C}(\text{O})\text{O}-$ ,  
 10  $-\text{C}(\text{O})(\text{CH}_2)_n-$  or a bond;

$n$  is an integer from 1 to 50;

each  $X_6$  and  $X_6$  is, independently, a bond, hydrogen or a hydrocarbyl group selected from  $\text{C}_1$ - $\text{C}_{10}$  alkyl,  $\text{C}_2$ - $\text{C}_{10}$  alkenyl,  $\text{C}_2$ - $\text{C}_{20}$  alkynyl,  $\text{C}_6$ - $\text{C}_{14}$  aryl,  $\text{C}_6$ - $\text{C}_{14}$  aralkyl,  $\text{C}_3$ - $\text{C}_{14}$  cycloalkyl,  $\text{C}_5$ - $\text{C}_{14}$  fused cycloalkyl,  $\text{C}_4$ - $\text{C}_{14}$  heterocycle,  $\text{C}_4$ - $\text{C}_{14}$  heterocyclalkyl,  $\text{C}_4$ - $\text{C}_{14}$  heteroaryl and  $\text{C}_4$ - $\text{C}_{14}$  heteroarylalkyl; wherein said hydrocarbyl group is substituted with at least two hydroxyl groups, and is optionally substituted with oxo, acyl, alkoxy, alkoxy carbonyl, alkyl, alkenyl, alkynyl, amino, amido, azido, aryl, heteroaryl, carboxylic acid, cyano, guanidino, halo, haloalkyl, haloalkoxy, hydrazino, ODMT, alkylsulfonyl, nitro, sulfide, disulfide, sulfone, sulfonate, sulfonamide, thiol, and thioalkoxy; provided that  $X_6$  is not hydrogen and  $X_6$  is  
 20 not a bond;

$\text{R}_1$  is hydrogen or a hydroxyl protecting group;

$\text{R}_4$  is a hydroxyl group or a protected hydroxyl group;

each  $\text{R}_5$  and  $\text{R}_{40}$  is, independently, hydrogen,  $\text{C}_1$ - $\text{C}_{10}$  alkyl,  $\text{C}_2$ - $\text{C}_{10}$  alkenyl,  $\text{C}_2$ - $\text{C}_{20}$  alkynyl,  $\text{C}_6$ - $\text{C}_{14}$  aryl or an amino-protecting group  
 30

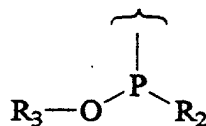
$\text{R}_{50}$  is hydrogen,  $\text{C}_1$ - $\text{C}_{10}$  alkyl,  $\text{C}_2$ - $\text{C}_{10}$  alkenyl,  $\text{C}_2$ - $\text{C}_{20}$  alkynyl,  $\text{C}_6$ - $\text{C}_{14}$  aryl,  $\text{C}_6$ - $\text{C}_{14}$  aralkyl,  $\text{C}_3$ - $\text{C}_{14}$  cycloalkyl, formyl,



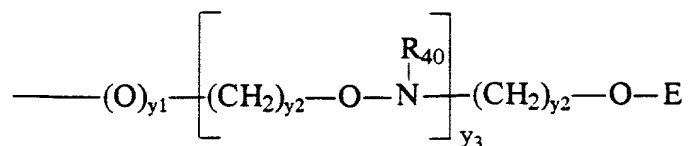
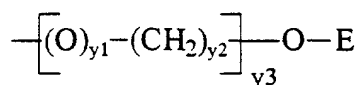
aminoalkyl or hydroxymethyl;

$R_6$  is hydrogen or an amino protecting group;

$R_{20}$  is hydrogen or a group of formula:



- 5  $R_2$  is  $-N(R_7)_2$ , or a heterocycloalkyl or heterocycloalkenyl ring containing from 4 to 7 atoms, and having up to 3 heteroatoms selected from nitrogen, sulfur, and oxygen;
- $R_7$  is straight or branched chain alkyl having from 1 to 10, 10 carbons;
- $R_3$  is a phosphorus protecting group;
- $R_{21}$  is hydrogen, hydroxyl, fluoro or a group of formula  $Z-R_{22}-(R_{23})_v$ ;
- $Z$  is O, S, NH or  $N-R_{22}-(R_{23})_v$ ;
- 15  $R_{22}$  is  $C_1-C_{20}$  alkyl,  $C_2-C_{20}$  alkenyl, or  $C_2-C_{20}$  alkynyl;
- $R_{23}$  is hydrogen, amino, halogen, hydroxyl, thiol, ketò, carboxyl, nitro, nitroso, nitrile, trifluoromethyl, trifluoromethoxy, O-alkyl, S-alkyl, NH-alkyl, N-dialkyl, O-aryl, S-aryl, NH-aryl, O-aralkyl, S-aralkyl, NH-aralkyl,
- 20 amino, N-phthalimido, imidazole, azido, hydrazino, hydroxylamino, isocyanato, sulfoxide, sulfone, sulfide, disulfide, silyl, aryl, heterocycle, carbocycle, intercalator, reporter molecule, conjugate, polyamine, polyamide, polyalkylene glycol, polyether, a group that enhances the
- 25 pharmacodynamic properties of oligonucleotides, or a group that enhances the pharmacokinetic properties of oligonucleotides;
- or  $R_{21}$  has one of the formulas:



wherein:

y<sub>1</sub> is 0 or 1;

each y<sub>2</sub> is, independently, 0 to 10;

5 y<sub>3</sub> is 1 to 10;

E is N(R<sub>41</sub>) (R<sub>42</sub>) or N=C(R<sub>41</sub>) (R<sub>42</sub>);

each R<sub>41</sub> and each R<sub>42</sub> is independently H, C<sub>1</sub>-C<sub>10</sub> alkyl, a  
nitrogen protecting group, or R<sub>41</sub> and R<sub>42</sub> taken together form a  
nitrogen protecting group; or R<sub>41</sub> and R<sub>42</sub> taken together with  
10 the N or C atom to which they are attached form a ring  
structure that can include at least one heterotom selected  
from N and O;

B is a nucleobase;

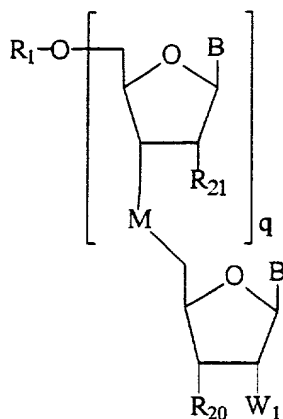
M is an optionally protected internucleoside linkage;

15 q is from zero to about 50; and

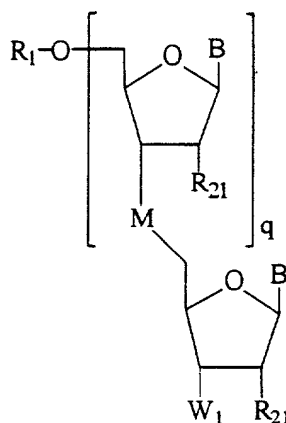
v is from zero to about 10.

In other preferred embodiments, the methods comprise  
the steps of:

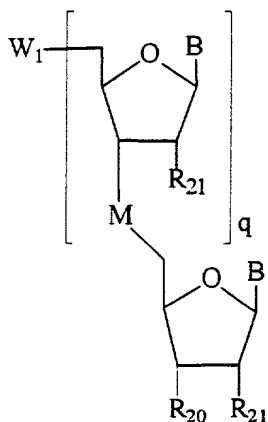
(a) providing a compound of formula IA, IB, IC or ID:



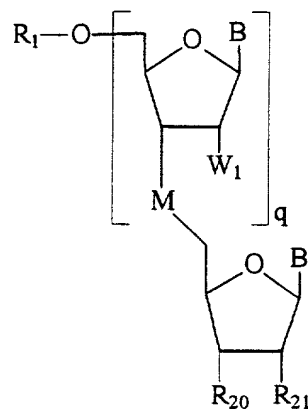
IA



IB



IC



ID

wherein:

$W_1$  is a linking group;

$R_1$  is H or a hydroxyl protecting group;

5  $B$  is a nucleobase;

each  $R_{21}$  is H, OH, F, or a group of formula  $Z-R_{22}-(R_{23})_v$ ;

$Z$  is O, S, NH, or  $N-R_{22}-(R_{23})_v$

$R_{22}$  is  $C_1-C_{20}$  alkyl,  $C_2-C_{20}$  alkenyl, or  $C_2-C_{20}$

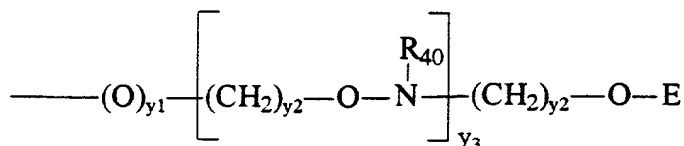
alkynyl;

10  $R_{23}$  is hydrogen, amino, halogen, hydroxyl, thiol, keto, carboxyl, nitro, nitroso, nitrile, trifluoromethyl, trifluoromethoxy, O-alkyl, S-alkyl, NH-alkyl, N-dialkyl, O-aryl, S-aryl, NH-aryl, O-aralkyl, S-aralkyl, NH-aralkyl, amino, N-phthalimido, imidazole, azido, hydrazino,

hydroxylamino, isocyanato, sulfoxide, sulfone, sulfide, disulfide, silyl, aryl, heterocycle, carbocycle, intercalator, reporter molecule, conjugate, polyamine, polyamide, polyalkylene glycol, polyether, a group that enhances the pharmacodynamic properties of oligonucleotides, or a group that enhances the pharmacokinetic properties of oligonucleotides;

v is from 0 to about 10;

or R<sub>21</sub> has one of the formulas:



wherein:

y<sub>1</sub> is 0 or 1;

each y<sub>2</sub> is, independently, 0 to 10;

y<sub>3</sub> is 1 to 10;

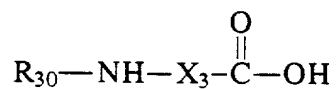
15 E is N(R<sub>41</sub>)(R<sub>42</sub>) or N=C(R<sub>41</sub>)(R<sub>42</sub>);

each R<sub>40</sub>, R<sub>41</sub> and R<sub>42</sub> is independently H, C<sub>1</sub>-C<sub>10</sub> alkyl, a nitrogen protecting group, or R<sub>41</sub> and R<sub>42</sub> taken together form a nitrogen protecting group; or R<sub>41</sub> and R<sub>42</sub> taken together with the N or C atom to which they are attached form a ring structure that can include at least one heterotom selected from N and O;

q is from 0 to about 50;

M is an optionally protected internucleoside linkage;

(b) reacting said compound of formula I with a  
25 compound of formula II:

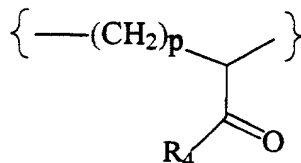


II

wherein:

$R_{30}$  is an amino protecting group;

5  $X_3$  is  $-CH(Z_1)-$  or a group of formula XI:

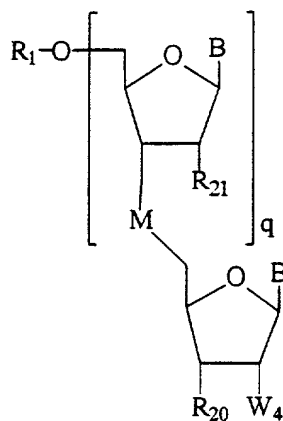


XI

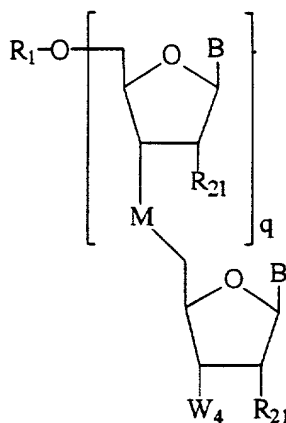
$Z_1$  is the sidechain of a naturally occurring amino acid, or a protected sidechain of a naturally occurring amino acid, wherein the amino acid is preferably glutamic acid;

10  $p$  is 1 or 2, with 2 being preferred;

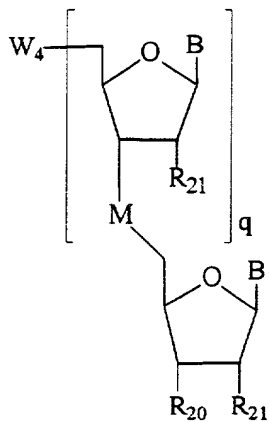
to form a compound of formula IVA, IVB, IVC, or IVD:



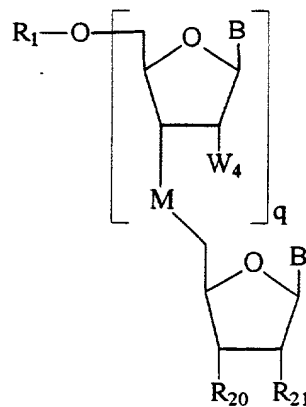
IVA



IVB



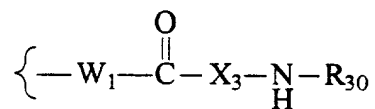
IVC



IVD

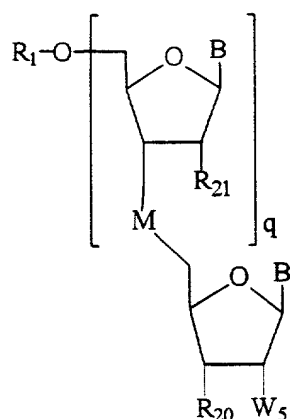
wherein:

$W_4$  has the formula:

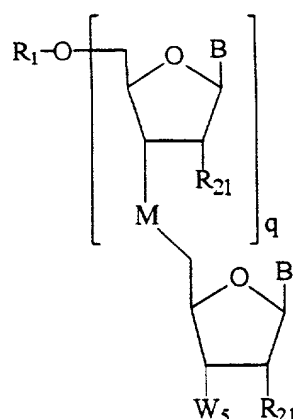


5 and

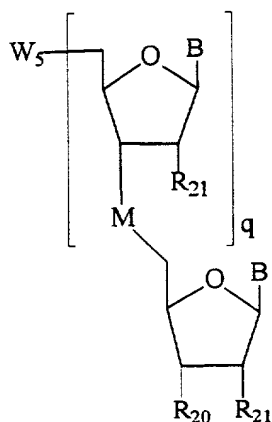
treating said compound of formula IVA, IVB, IVC or IVD with a deprotecting reagent to form a compound of formula VA, VB, VC or VD:



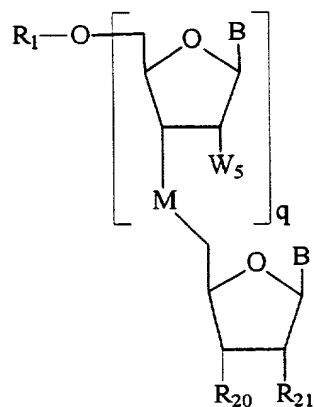
VA



VB

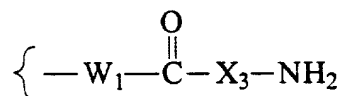


VC



VD

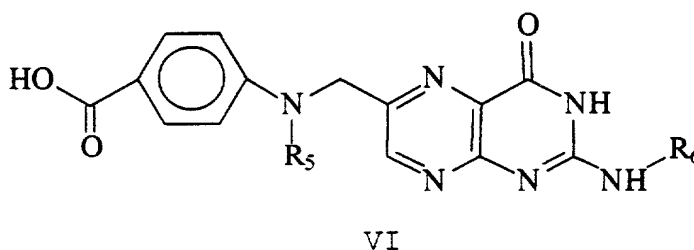
wherein  $W_5$  has the formula:



As used herein, the term "reacting" means placing  
 5 moieties together in a container under conditions of time,  
 temperature and pressure such that the indicated reaction  
 occurs. As used herein, the term "treating said compound of  
 formula IVA, IVB, IVC or IVD with a deprotecting reagent to  
 form a compound of formula VA, VB, VC or VD" means causing  
 10 the formation of a compound of formula VA, VB, VC or VD by  
 effecting the deprotection of a compound of formula IVA, IVB,

IVC or IVD using a deprotection reagent. The term "deprotecting reagent" denotes a reagent that is effective to remove the protecting group  $R_{30}$ . Suitable deprotecting reagents can be found in Greene and Wuts, *Protective Groups in Organic Synthesis*, Chapter 2, 2d ed, John Wiley & Sons, New York, 1991, and *Oligonucleotides And Analogues A Practical Approach*, Ekstein, F. Ed., IRL Press, N.Y, 1991, each of which are incorporated by reference herein in their entirety. Preferred deprotection reagents include, but are not limited to, 0.1 M DBU in acetonitrile or a 10% solution of piperidine in acetonitrile.

In some preferred embodiments, the methods of the invention further comprise condensing said compound of formula V with a compound of formula VI:



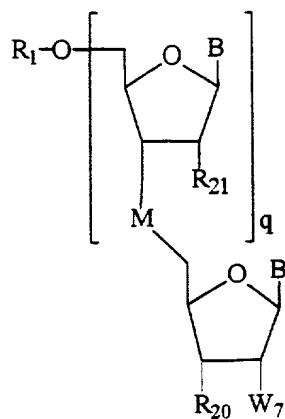
wherein:

$R_5$  is H or an amino protecting group;

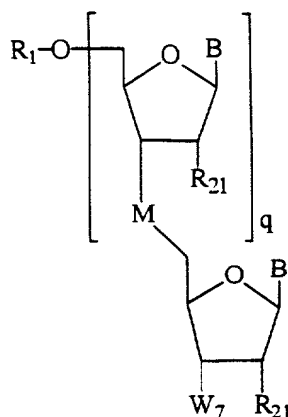
20  $R_6$  is H or an amino protecting group;

to form a compound of formula VIIA, VIIB, VIIC, or VIID:

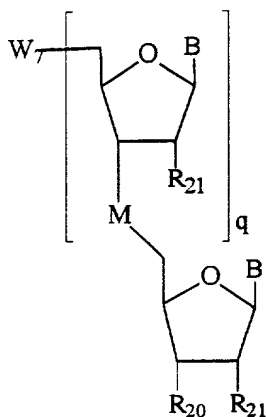




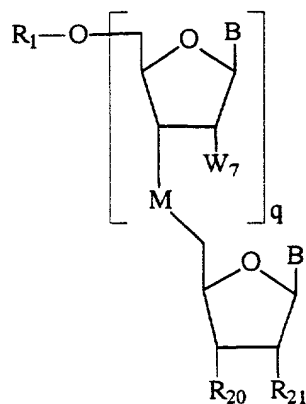
VIIA



VIIB

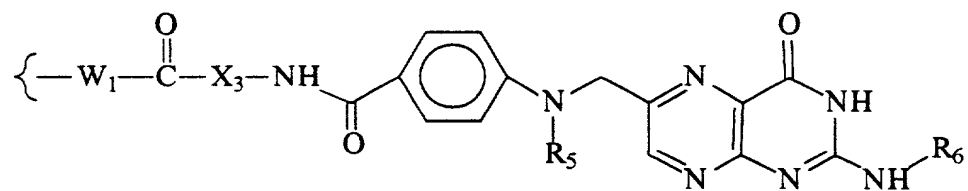


VIIC



VIID

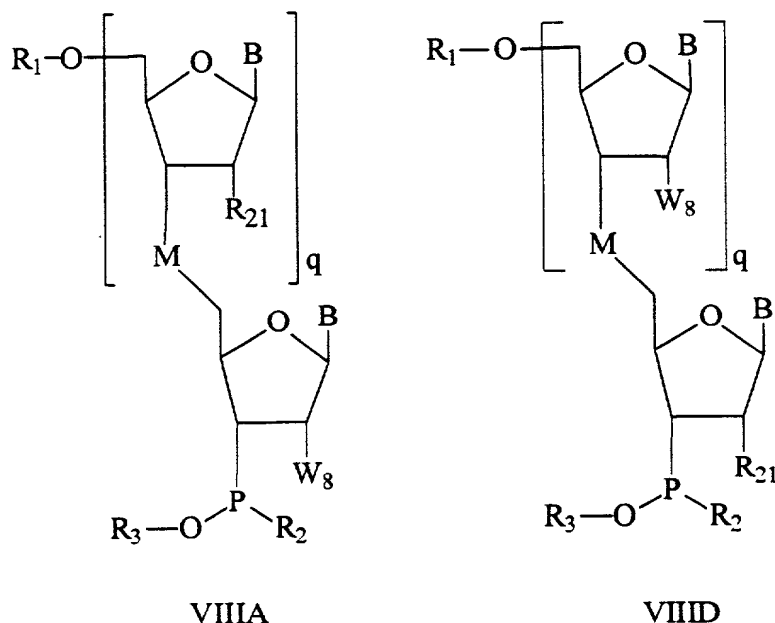
wherein  $W_7$  has the formula:



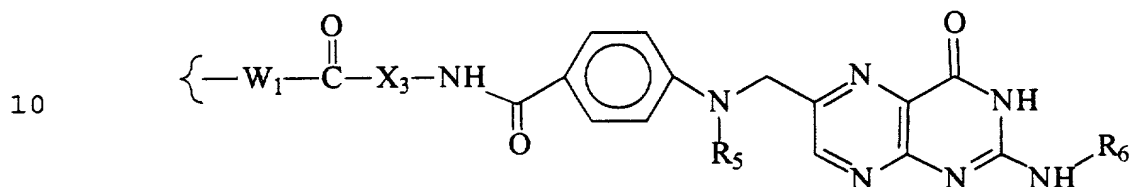
- 5 As used herein the term "condensing said compound of formula V with a compound of formula VI" means placing a compound of formula V together with a compound of formula VIA-D under conditions of time, temperature and pressure effective to form a condensation product of formula VIIA-D.
- 10 The condensation is preferably performed in the presence of

dicyclohexylcarbodiimide (DCC) and HONB as activating reagent in anhydrous DMF. Other suitable reagents for this purpose include HATU, HBTU and PyBrop.

Some preferred embodiments of the methods of the invention further comprise contacting said compound of formula VIIA or VIID with a phosphitylating reagent to form a compound of formula VIIIA or VIIIA-D:

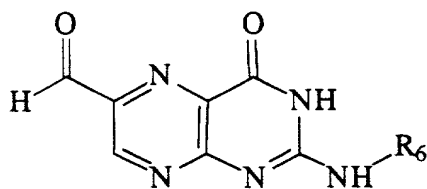


wherein  $W_7$  has the formula:



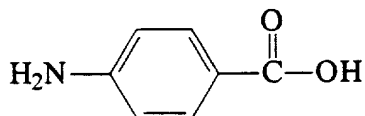
The phosphitylated compound of formula VIIIA or VIIID can be used as a synthon in standard oligonucleotide synthetic regimes, as disclosed in, for example, Ekstein, *infra*.

15 In some preferred embodiments, compounds of formula VI are prepared by reacting a compound of formula IX:



IX

with a compound of formula X:

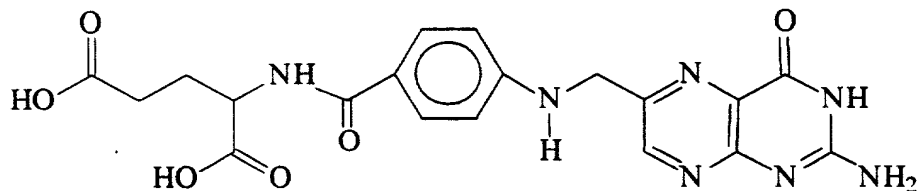


X

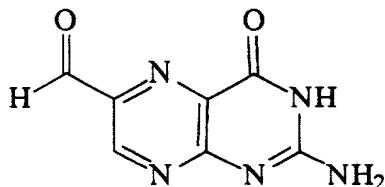
5

and treating the product of said reaction with a protecting group reagent to form said compound of formula VI.

In accordance with preferred embodiments of the methods of the invention, folic acid conjugates and/or  
 10 derivatives can be prepared from pterin aldehyde, which is obtained from folic acid. Preferably, the folic acid from which the pterin aldehyde is obtained is isolated from a natural source, or purchased commercially. Thus, in some especially preferred embodiments, compounds of formula IX are  
 15 prepared by reacting folic acid:



with a reagent effective to form pterin aldehyde:



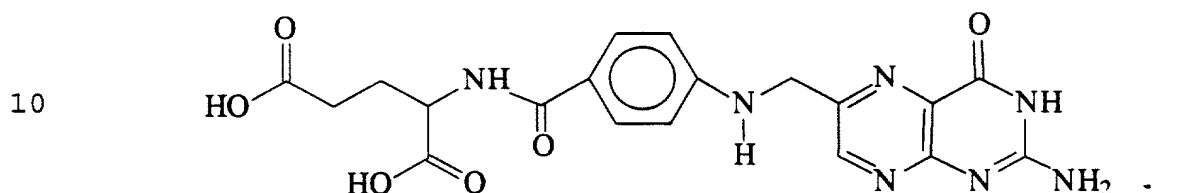
and

20 protecting the amino group of said pterin aldehyde.

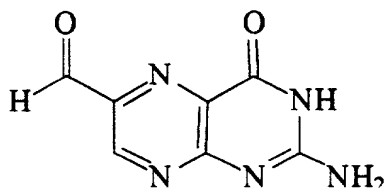
In preferred embodiments, the pterin aldehyde is obtained from the reaction of folic acid and HBr/Br<sub>2</sub>,

preferably at a temperature of 95-100°C. The amino group of the pterin aldehyde can be selectively protected by, for example, treatment with isobutyric anhydride at a temperature of 130° C for 30 hours. The 4-oxo moiety of the pterins may also be protected as the diphenylcarbamate species by using, for example, diphenylcarbamoyl chloride.

Also provided in accordance with the present invention are methods for the preparation of a folic acid derivative comprising the steps of reacting folic acid:

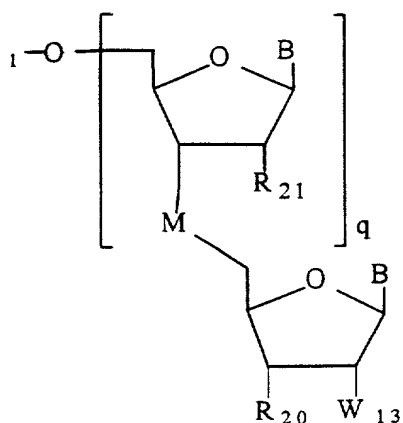


with a reagent effective to form pterin aldehyde:

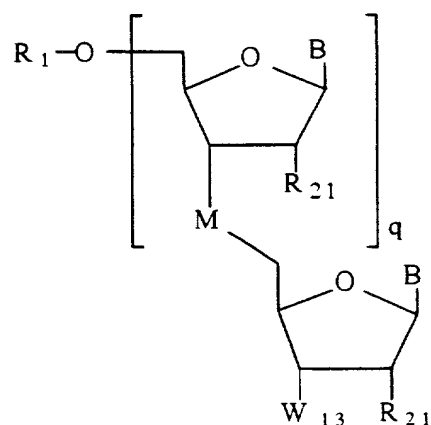


and preferably protecting the amino group of said pterin aldehyde.

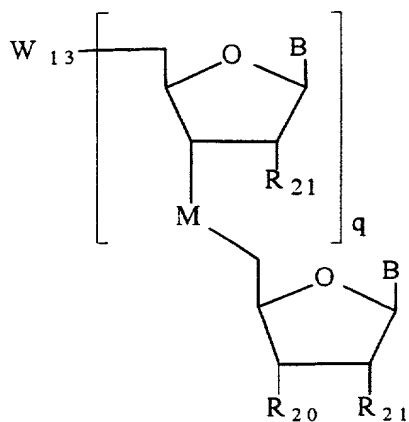
15 The present invention also provides compounds having the formula XIIIA, XIIB, XIIIC or XIIID:



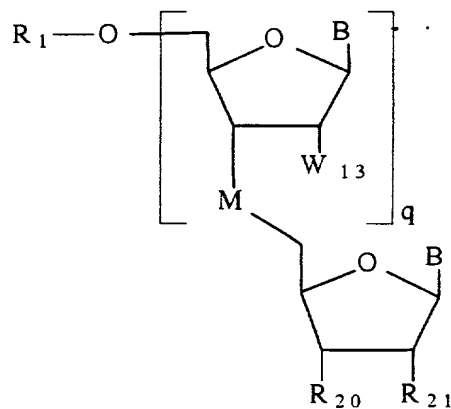
XIII A



XIII B



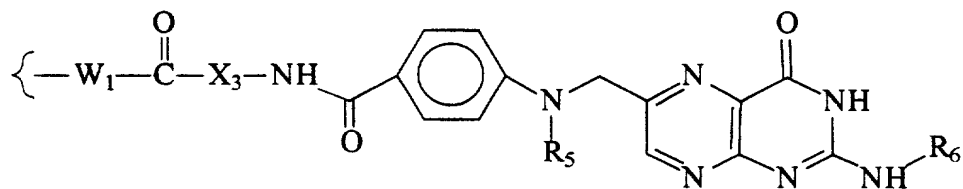
XIII C



XIII D

wherein:

$W_{13}$  has the formula:



5

$R_1$  is H or a hydroxyl protecting group;

B is a nucleobase;

each  $R_{21}$  is H, OH, F, or a group of formula  $Z-R_{22}-(R_{23})_v$ ;

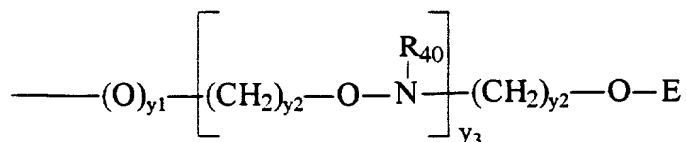
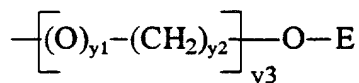
Z is O, S, NH or  $N-R_{22}-(R_{23})_v$ ;

$R_{22}$  is  $C_1-C_{20}$  alkyl,  $C_2-C_{20}$  alkenyl, or  $C_2-C_{20}$

10 alkynyl;

$R_{23}$  is hydrogen, amino, halogen, hydroxyl, thiol, keto, carboxyl, nitro, nitroso, nitrile, trifluoromethyl, trifluoromethoxy, O-alkyl, S-alkyl, NH-alkyl, N-dialkyl, O-aryl, S-aryl, NH-aryl, O-aralkyl, S-aralkyl, NH-aralkyl,  
 5 amino, N-phthalimido, imidazole, azido, hydrazino, hydroxylamino, isocyanato, sulfoxide, sulfone, sulfide, disulfide, silyl, aryl, heterocycle, carbocycle, intercalator, reporter molecule, conjugate, polyamine, polyamide, polyalkylene glycol, polyether, a group that enhances the  
 10 pharmacodynamic properties of oligonucleotides, or a group that enhances the pharmacokinetic properties of oligonucleotides;

or  $R_{21}$  has one of the formulas:



15 wherein:

$y_1$  is 0 or 1;

each  $y_2$  is, independently, 0 to 10;

$y_3$  is 1 to 10;

E is  $N(R_{41})(R_{42})$  or  $N=C(R_{41})(R_{42})$ ;

20 each  $R_{40}$ ,  $R_{41}$  and  $R_{42}$  is independently H,  $C_1$ - $C_{10}$  alkyl, a nitrogen protecting group, or  $R_{41}$  and  $R_{42}$  taken together form a nitrogen protecting group; or  $R_{41}$  and  $R_{42}$  taken together with the N or C atom to which they are attached form a ring structure that can include at least one  
 25 heterotom selected from N and O;

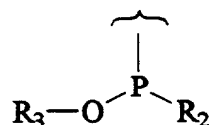
$v$  is from 0 to about 10;

$q$  is from 0 to about 50;

M is an optionally protected internucleoside linkage;

$W_1$  is a linking group, O, NH or S, with a linking group being preferred;

$R_{20}$  is H or a group of formula:



5  $R_2$  is  $-N(R_7)_2$ , or a heterocycloalkyl or heterocycloalkenyl ring containing from 4 to 7 atoms, and having up to 3 heteroatoms selected from nitrogen, sulfur, and oxygen;

10  $R_7$  is straight or branched chain alkyl having from 1 to 10 carbons;

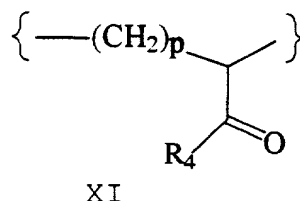
$R_3$  is a phosphorus protecting group;

$n$  is from 1 to about 10;

$R_5$  is H or an amino protecting group;

$R_6$  is H or an amino protecting group;

15  $X_3$  is  $-CH(Z_1)-$  or a group of formula XI:

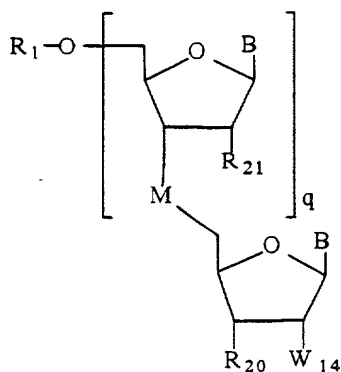


20  $Z_1$  is the sidechain of a naturally occurring amino acid, or a protected sidechain of a naturally occurring amino acid, wherein the amino acid is preferably glutamic acid;

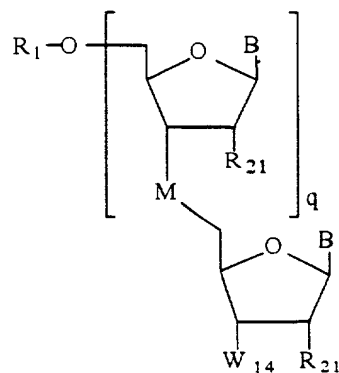
$p$  is 1 or 2, with 2 being preferred; and

$R_4$  is a hydroxyl group, or a protected hydroxyl group.

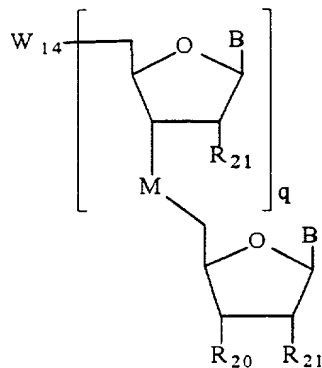
The present invention also provides compositions comprising a compound of formula XIIIA-D wherein  $X_3$  has the  
25 formula XI and  $p$  is 2, said composition being substantially free of a compound of formula XIVA, XIVB, XIVC, or XIVD:



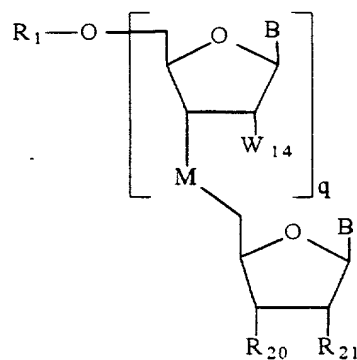
XIV A



XIV B



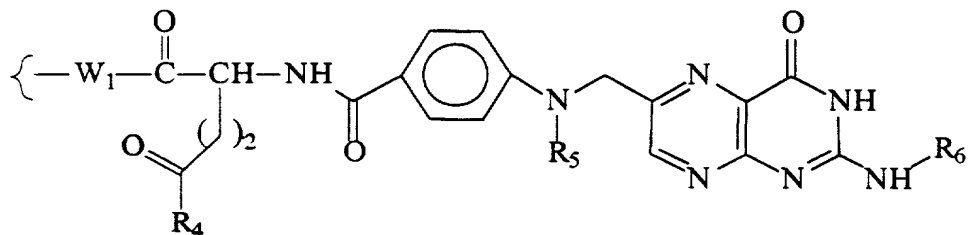
XIV C



XIV D

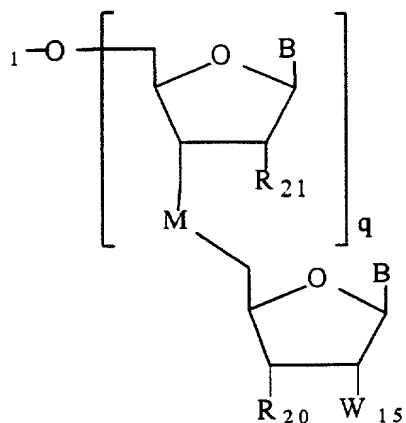
wherein:

$W_{14}$  has the formula:

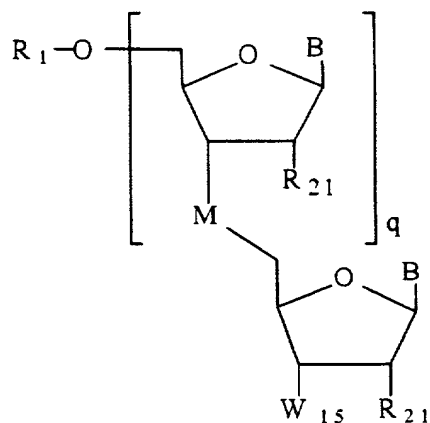


- 5 Also provided in accordance with the present invention are compositions comprising a compound of formula XIII A-D wherein  $X_3$  has the formula XII and  $m$  is 2, said composition being substantially free of a compound of formula XVA, XVB, XVC or XVD:

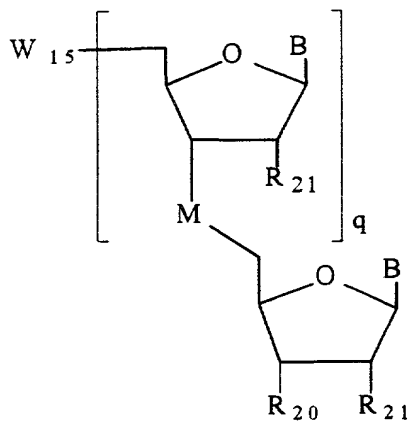




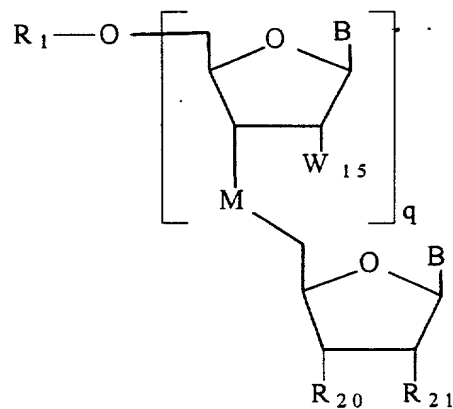
XVA



XVB

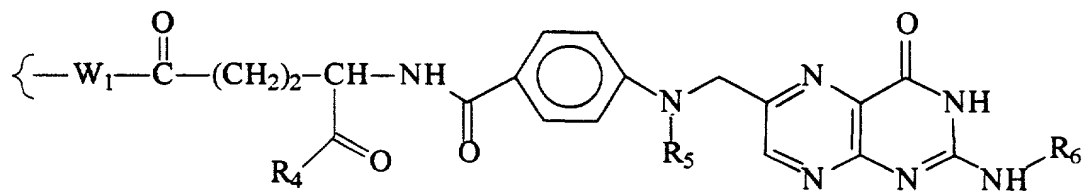


XVC

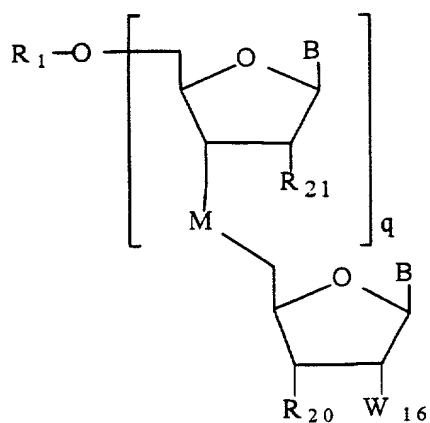


XVD

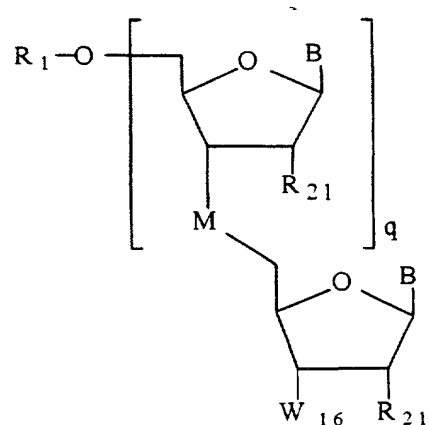
wherein  $W_{15}$  has the formula:



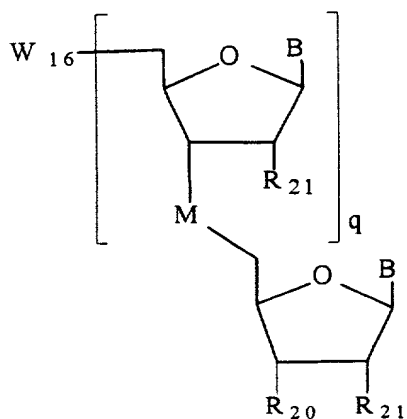
The present invention further provides compounds  
5 having the formula XVIA, XVIB, XVIC or XVID:



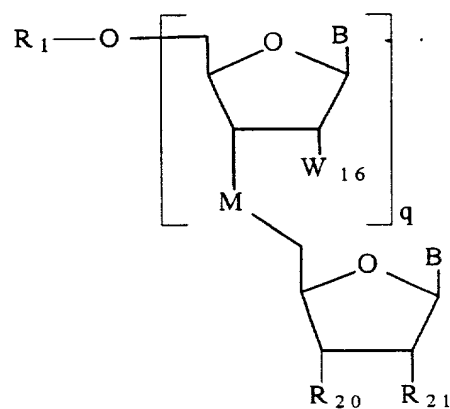
XVIIA



XVIIB



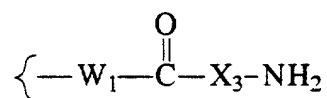
XVIIIC



XVIIID

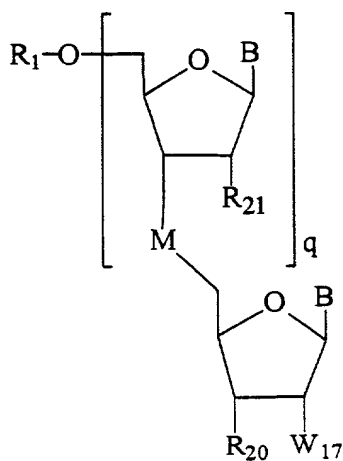
wherein:

$W_{16}$  has the formula:

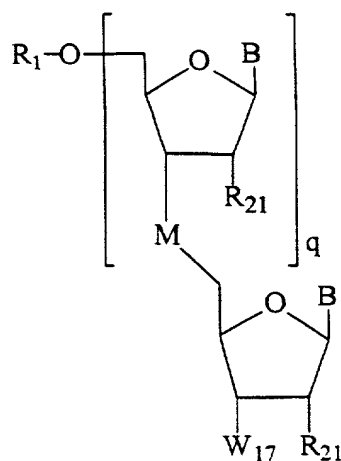


5 wherein the constituent variables are as defined above.

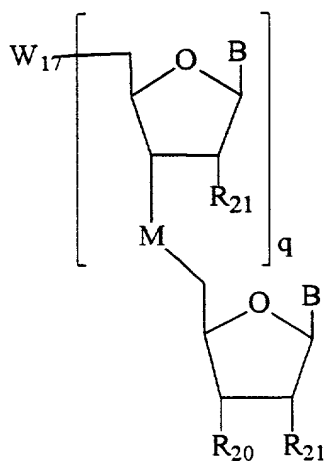
Also provided in accordance with the present invention are compounds having the formula XVIIIA, XVIIIB, XVIIIC or XVIIID:



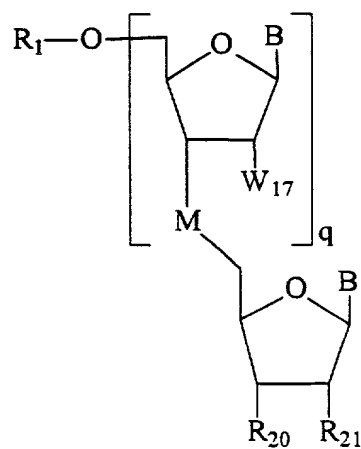
XVIIA



XVIIIB



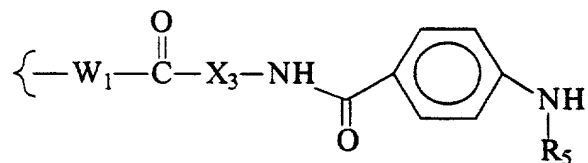
XVIIIC



XVIIID

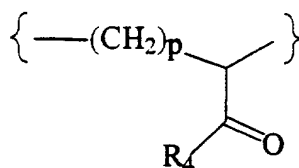
wherein:

$W_{17}$  has the formula:



5 wherein the constituent variables are as defined above.

In some preferred embodiments of the foregoing methods and compounds,  $X_3$  is a group of formula XI:

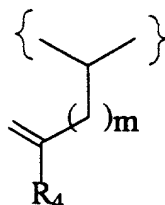


XI

wherein:

p is 1 or 2, with 2 being preferred;

- 5       $\text{R}_4$  is a hydroxyl group, or a protected hydroxy group;  
or  $\text{X}_3$  is a group of formula XII:



XII

wherein m is 1 or 2, with 2 being preferred.

- 10      In some preferred embodiments of the foregoing compounds and methods, q is 0.

In further preferred embodiments of the foregoing compounds and methods,  $\text{X}_3$  is a group of formula XI, preferably wherein p is 2.

- 15      In further preferred embodiments of the foregoing compounds and methods,  $\text{X}_3$  is a group of formula XII, preferably wherein m is 2.

- In preferred embodiments of the foregoing compounds and methods,  $\text{W}_1$  has the formula  $-\text{O}-(\text{CH}_2)_n-\text{NH}-$ , where n is  
20 from 1 to about 10, with 6 being preferred.

- In some preferred embodiments of the foregoing compounds and methods,  $\text{R}_1$  is dimethoxytrityl. In further preferred embodiments,  $\text{R}_4$  is *t*-butoxy. In still further preferred embodiments,  $\text{R}_5$  is trifluoroacetyl. In further  
25 preferred embodiments,  $\text{R}_6$  is  $-\text{C}(=\text{O})-\text{CH}(\text{CH}_3)_2$ . In yet further preferred embodiments,  $\text{R}_{30}$  is fluorene-9-yl methoxycarbonyl.

In some preferred embodiments of the foregoing compounds and methods,  $\text{R}_1$  is dimethoxytrityl,  $\text{W}_1$  has the

formula  $-O-(CH_2)_n-NH-$  where  $n$  is 6,  $X_3$  has the formula XI where  $p$  is 2,  $R_4$  is *t*-butoxy,  $R_5$  is trifluoroacetyl,  $R_6$  is  $-C(=O)-CH(CH_3)_2$ , and  $R_{30}$  is FMOC.

In further preferred embodiments,  $R_1$  is  
5 dimethoxytrityl,  $W_1$  has the formula  $-O-(CH_2)_n-NH-$  where  $n$  is 6,  $X_3$  has the formula XII where  $m$  is 2,  $R_4$  is *t*-butoxy,  $R_5$  is trifluoroacetyl,  $R_6$  is  $-C(=O)-CH(CH_3)_2$ , and  $R_{30}$  is FMOX.

In preferred embodiments, the conjugates of the present invention include folate moieties such as, but not  
10 limited to, folic acid, folic acid derivatives and analogs, antifolates and deazafolates. As used herein, the term "folate" shall include all such structures. Examples of such folates include folic acid, dihydrofolic acid,  
15 tetrahydrofolic acid, folinic acid, pteropolyglutamic acid, dihydrofolates, tetrahydrofolates, tetrahydropterins, 1-deaza, 3-deaza, 5-deaza, 8-deaza, 10-deaza, 1,5-dideaza, 5,10-dideaza, 8,10-dideaza and 5,8-dideaza folate analogs, and antifolates. In some preferred embodiments, folic acid is employed as the folate moiety in the compounds and methods  
20 of the invention.

The folate moiety or moieties may be attached to one or more sites on a mononucleoside, mononucleotide  
phosphoramidite, or oligonucleotide of the invention. These sites include one or more of, but are not limited to, the 2'-  
25 position, 3'-position, 5'-position, the internucleotide linkage, and a nucleobase atom of any nucleotide residue. The attachment of folates to such structures can be performed, according to some preferred embodiments of the invention, using a linking group, or without the use of such  
30 a linking group.

In some preferred embodiments of the present invention, folate moiety or moieties may be attached or covalently linked to nucleosidic or non-nucleosidic molecules or groups, to form folate conjugates. The folate moiety may  
35 optionally be connected to a nucleosidic or non-nucleosidic

molecule via an amino acid. Preferred non-nucleosidic groups include hydrocarbyl groups. By the term "hydrocarbyl" is meant a group or moiety comprising carbon and hydrogen atoms. Accordingly, hydrocarbyl includes straight and branched  
5 alkyl, alkenyl and alkynyl groups, as well as cycloalkyl, fused cycloalkyl, aryl, aralkyl, heterocyclyl, heteroaryl, heterocyclylalkyl and heteroarylalkyl. The hydrocarbyl group may be substituted with substituents including, but not limited to, oxo, acyl, alkoxy, alkoxycarbonyl, alkyl,  
10 alkenyl, alkynyl, amino, amido, azido, aryl, heteroaryl, carboxylic acid, cyano, guanidino, halo, haloalkyl, haloalkoxy, hydrazino, ODMT, alkylsulfonyl, nitro, sulfide, sulfone, sulfonamide, thiol and thioalkoxy.

By "hetero" is meant a group incorporating one or more  
15 heteroatoms, such as N, O, S, SO and SO<sub>2</sub>. "Cyclo" or "cyclyl" includes a cyclic group which may be mono-, bi- or tricyclic, and may be substituted with substituents such as oxo, acyl, alkoxy, alkoxycarbonyl, alkyl, alkenyl, alkynyl, amino, amido, azido, aryl, heteroaryl, carboxylic acid,  
20 cyano, guanidino, halo, haloalkyl, haloalkoxy, hydrazino, ODMT, alkylsulfonyl, nitro, sulfide, sulfone, sulfonamide, thiol and thioalkoxy.

In some preferred embodiments of the compounds and methods of the present invention, one or more folate moieties  
25 is attached to a nucleoside or oligonucleotide portion of the conjugate via linking groups, to form an oligonucleotide-folate conjugate. The oligonucleotide of the oligonucleotide-folate conjugate may optionally be connected to a folate moiety via an amino acid. Preferred amino acids  
30 include naturally- and non-naturally-occurring amino acids. It is more preferred that the amino acids be naturally-occurring amino acids.

Preferred linking groups of the invention include, but are not limited to, w-aminoalkoxy linkers, w-aminoalkylamino  
35 linkers, cysteamine, heterobifunctional linkers,

homobifunctional linkers, and a universal linker (derived from 3-dimethoxytrityloxy-2-aminopropanol). A particularly preferred linking group for the synthesis of oligonucleotide-folate conjugates of the invention is a 6-aminohexyloxy group.

The oligonucleotide-folate conjugates of the present invention contain one or more optionally protected oligonucleoside linkages, represented by variable "M" in the formulas provided herein. Representative internucleoside linkages include optionally protected phosphite, phosphodiester, phosphorothioate, phosphorodithioate, and alkyl phosphonate linkages.

As used herein, the term "protected" means that the indicated moiety has appended a protecting group thereto. In some preferred embodiments of the invention compounds contain one or more protecting groups. A wide variety of protecting groups can be employed in the methods of the invention. In general, protecting groups render chemical functionalities inert to specific reaction conditions, and can be appended to and removed from such functionalities in a molecule without substantially damaging the remainder of the molecule.

Representative hydroxyl protecting groups, for example, are disclosed by Beaucage, et al., *Tetrahedron* **1992**, *48*, 2223-2311. Further hydroxyl protecting groups, as well as other representative protecting groups are disclosed in Greene and Wuts, *Protective Groups in Organic Synthesis*, Chapter 2, 2d ed, John Wiley & Sons, New York, 1991, and *Oligonucleotides And Analogues A Practical Approach*, Ekstein, F. Ed., IRL Press, N.Y, 1991, each of which is hereby incorporated by reference in its entirety.

Examples of hydroxyl protecting groups include, but are not limited to, t-butyl, t-butoxymethyl, methoxymethyl, tetrahydropyranyl, 1-ethoxyethyl, 1-(2-chloroethoxy)ethyl, 2-trimethylsilylethyl, p-chlorophenyl, 2,4-dinitrophenyl, benzyl, 2,6-dichlorobenzyl, diphenylmethyl, p,p'-

dimnitrobenzhydryl, p-nitrobenzyl, triphenylmethyl, trimethylsilyl, triethylsilyl, t-butyldimethylsilyl, t-butyldiphenylsilyl, triphenylsilyl, benzoylformate, acetate, chloroacetate, trichloroacetate, trifluoroacetate, pivaloate, benzoate, p-phenylbenzoate, 9-fluorenylmethyl carbonate, mesylate and tosylate.

Amino-protecting groups stable to acid treatment are selectively removed with base treatment, and are used to make reactive amino groups selectively available for substitution. Examples of such groups are the Fmoc (E. Atherton and R.C. Sheppard in *The Peptides*, S. Udenfriend, J. Meienhofer, Eds., Academic Press, Orlando, 1987, volume 9, p.1), and various substituted sulfonylethyl carbamates exemplified by the Nsc group (Samukov et al., *Tetrahedron Lett*, 1994, 35:7821; Verhart and Tesser, *Rec. Trav. Chim. Pays-Bas*, 1987, 107:621).

Additional amino-protecting groups include but are not limited to, carbamate-protecting groups, such as 2-trimethylsilylethoxycarbonyl (Teoc), 1-methyl-1-(4-biphenyl)ethoxycarbonyl (Bpoc), t-butoxycarbonyl (BOC), allyloxycarbonyl (Alloc), 9-fluorenylmethyloxycarbonyl (Fmoc), and benzyloxycarbonyl (Cbz); amide-protecting groups, such as formyl, acetyl, trihaloacetyl, benzoyl, and nitrophenylacetyl; sulfonamide-protecting groups, such as 2-nitrobenzenesulfonyl; and imine- and cyclic imide-protecting groups, such as phthalimido and dithiasuccinoyl. Equivalents of these amino-protecting groups are also encompassed by the compounds and methods of the present invention.

In some especially preferred embodiments, the nucleoside components of the oligonucleotide-folate conjugates are connected to each other by optionally protected phosphorothioate internucleoside linkages. Representative protecting groups for phosphorus containing internucleoside linkages such as phosphphite, phosphodiester and phosphorothioate linages include  $\beta$ -cyanoethyl,



diphenylsilylethyl,  $\delta$ -cyanobutenyl, cyano *p*-xylyl (CPX), N-methyl-N-trifluoroacetyl ethyl (META), acetoxymethoxy ethyl (APE) and butene-4-yl groups. See for example U.S. Patents Nos. 4,725,677 and Re. 34,069 ( $\beta$ -cyanoethyl); Beaucage, S.L. and Iyer, R.P., *Tetrahedron*, **49** No. 10, pp. 1925-1963 (1993); Beaucage, S.L. and Iyer, R.P., *Tetrahedron*, **49** No. 46, pp. 10441-10488 (1993); Beaucage, S.L. and Iyer, R.P., *Tetrahedron*, **48** No. 12, pp. 2223-2311 (1992).

The present invention provides compounds that have a plurality of linked nucleosides and where at least one of the nucleosides is a 2'-functionalized nucleoside having a folate molecule linked to the 2'-position of the nucleoside; a heterocyclic base functionalized nucleoside having a folate molecule linked to the heterocyclic base of the nucleoside, a 5' terminal nucleoside having a folate molecule linked to the 5'-position of the nucleoside, a 3' terminal nucleoside having a folate molecule linked to the 3'-position of the nucleoside, or an inter-strand nucleoside having a folate molecule linked to an inter-strand linkage linking said inter-strand nucleoside to an adjacent nucleoside.

Oligonucleotide-folate conjugate compounds have been synthesized by the use of an oligonucleotide that bears a pendant reactive functionality such as that derived from the attachment of a linking molecule onto the oligonucleotide. This reactive oligonucleotide may be reacted directly with commercially available folates or folates that are synthesized bearing a variety of protecting groups. For example, reaction of folic acid with aminolinked oligonucleotides using carbodiimide activation has been reported by Low et al., U.S. Patent 5,108,921. However, this procedure is difficult and of limited utility. In contrast, the methods of the present invention facilitate the synthesis of oligonucleotide-folate conjugates by the use of, in some preferred embodiments, nucleoside monomers that have been appropriately conjugated with folates. Such folate-

nucleoside conjugates are prepared according to some preferred embodiments of the methods of the present invention via reaction of pteric acid derivatives with the desired 2'-aminohexyloxy-nucleoside bearing a glutamate on the pendant  
5 2'-amine. The glutamate may be attached to the nucleoside and linker via either its  $\alpha$ - or  $\gamma$ -carboxyl groups. Thus, the methods of the present invention advantageously provide for complete control over the regiochemistry and structure of the conjugate, thus providing a significant benefit over existing  
10 synthetic techniques.

It will be appreciated that through the choice of appropriate protecting groups and cleavage reactions, the methods of the present invention allow the regioselective synthesis of  $\alpha$ - and  $\gamma$ -isomers of folates conjugated to any  
15 desired nucleoside, preferably at the 2'-position of the nucleoside. In addition, conjugates of the invention can also be prepared wherein the folate or folate-linker moiety is attached at other locations on the nucleoside, as described, *supra*.

20 In some preferred embodiments, the folate-nucleoside conjugates of the present invention are converted to their corresponding 3'-phosphoramidites via phosphitylation reactions, as described herein. Methods for the use of such phosphoramidites in the synthesis of oligonucleotide-folate  
25 conjugates are disclosed herein, and in, for example, Ekstein, *supra*.

The present invention also provides methods of increasing cellular uptake of an oligonucleotide comprising contacting an organism with an oligonucleotide of the  
30 invention.

Oligonucleotide-folate conjugates of the present invention can be included in compositions that further include one more inert carrier compounds.

Antisense therapeutics can be practiced in a plethora  
35 of various organisms ranging from unicellular prokaryotic and

eukaryotic organisms to multicellular eukaryotic organisms. Any organism that utilizes DNA-RNA transcription or RNA-protein translation as a fundamental part of its hereditary, metabolic or cellular control is susceptible to antisense  
5 therapeutics and/or prophylactics. Seemingly diverse organisms such as bacteria, yeast, protozoa, algae, all plant and all higher animal forms, including warm-blooded animals, can be treated by antisense therapy. Further, since each of the cells of multicellular eukaryotes also includes both DNA-  
10 RNA transcription and RNA-protein translation as an integral part of its cellular activity, antisense therapeutics and/or diagnostics can also be practiced on such cellular populations. Furthermore, many of the organelles, e.g. mitochondria and chloroplasts, of eukaryotic cells also include  
15 transcription and translation mechanisms. As such, single cells, cellular populations or organelles can also be included within the definition of organisms that are capable of being treated with antisense therapeutics or diagnostics. As used herein, therapeutics is meant to include both the  
20 eradication of a disease state, killing of an organism, e.g. bacterial, protozoan or other infection, or control of erratic or harmful cellular growth or expression.

While we do not wish to be bound by any particular theory, it is believed that the presence of many nuclear  
25 proteins in the nucleus is due to their selective entry through the nuclear envelope rather than to their selective retention within the nucleus after entry. By this mechanism, the nucleus is able to selectively take up certain proteins and not others. The uptake is based upon the sequence of the  
30 peptide or protein, which provides a selective signal sequence that allows accumulation of the peptide or protein in the nucleus. One such peptide signal sequence is found as part of the SV40 large T-antigen. See: Dingwell and Laskey, *Ann. Rev. Cell Bio.*, 2:367 (1986); Yoneda et. al., "Synthetic  
35 Peptides Containing a Region of SV40 Large T-Antigen Involved

In Nuclear Localization Direct The Transport Of Proteins Into The Nucleus", *Experimental Cell Research*, 170:439 (1987); and Wychowski et. al., "The Intracellular Location Of Simian Virus 40 Polypeptides VP2 and VP3 Depends On a Specific Amino Acid Sequence" *J. Virol.*, 61:3862 (1986).

The present invention provides oligonucleotide-folate conjugates that have improved therapeutic properties, including improved transfer across cellular membranes, and also provides methods for the preparation of said oligonucleotide-folate conjugates. In accordance with the present invention there are provided oligonucleotide-folate conjugates and compounds that comprise a plurality of linked nucleosides wherein at least one of the nucleosides is functionalized with a folate molecule, preferably by linking said folate molecule to the nucleoside with a linking moiety.

In accordance with the present invention, a folate molecule is attached to at least one nucleoside in an antisense diagnostic or therapeutic agent to assist in the transfer of the antisense therapeutic or diagnostic agent across cellular membranes. Such antisense diagnostic or therapeutic agent is preferably a nucleic acid or oligonucleotide formed of a plurality of linked nucleosides of a sequence that are "antisense" to a region of an RNA or DNA of interest. Thus one or more nucleosides of the oligonucleotide are conjugated to include a folate molecule linked to the nucleoside with or without the intermediacy of a linking group. For the purposes of identification, such conjugated nucleosides can be characterized as folate bearing nucleosides or folate-nucleoside conjugates. The linked nucleosides having at least one conjugated nucleoside within their sequence demonstrate enhanced antisense activity when compared to like linked nucleoside or oligonucleotides of the same sequence that are not conjugated. The linked nucleosides having at least one functionalized nucleoside within their sequence further demonstrate increased transfer

across cellular membranes.

Particularly useful as folate molecules for the preparation of folate conjugates of the present invention are molecules selected from the group consisting of folic acid, dihydrofolic acid, tetrahydrofolic acid, folinic acid, pteropolyglutamic acid, dihydrofolates, tetrahydrofolates, tetrahydropterins, 1-deaza, 3-deaza, 5-deaza, 8-deaza, 10-deaza, 1,5-dideaza, 5,10-dideaza, 8,10-dideaza and 5,8-dideaza folate analogs. The terms deaza and dideaza analogs refers to the art recognized folate analogs that have a carbon atom substituted for one or two nitrogen atoms in the naturally occurring folic acid structure. The foregoing folic acid derivatives are conventionally referred to as folates because they have a structural resemblance to folic acid and because they bind to folate receptors. Folate molecules also include additional molecules that are known to bind to folate receptors and include, but are not limited to, antifolates such as aminopterin, methotrexate, N<sup>10</sup>-methylfolate, 2-desaminohydroxyfolate, 1-deazamethotrexate, 3-deazamethotrexate, and dichloromethotrexate.

Various linking groups are useful for linking folate molecules to nucleosides in accordance with the invention. Preferred linking groups of the invention include, but are not limited to, w-aminoalkoxy linkers, w-aminoalkylamino linkers, cysteamine, heterobifunctional linkers, homobifunctional linkers, and a universal linker (derived from 3-dimethoxytrityloxy-2-aminopropanol). a particularly preferred linking molecule for the synthesis of oligonucleotide-folate conjugates of the invention is the 6-aminohexyloxy group.

Linking groups such as w-aminoalkoxy moieties and w-aminoalkylamino moieties, are particularly useful for linking molecular groups such as steroid molecules or reporter molecules to, for example, the 2'-position of a nucleoside. a variety of heterobifunctional and homobifunc-

tional linking moieties are available from the Pierce Co. (Rockford, Il). Such heterobifunctional and homobifunctional linking moieties are particularly useful in conjunction with the w-aminoalkoxy and w-aminoalkylamino moieties to form

5 extended linkers useful for linking peptides, proteins and other molecules such as folates to a nucleoside. Further useful linking groups that are commercially available are 5'-Amino-Modifier C6 and 3'-Amino-Modifier reagents, both available from Glen Research Corporation (Sterling, VA). 5'-

10 Amino-Modifier C6 is also available from ABI (Applied Biosystems Inc., Foster City, CA) as Aminolink-2, while the 3'-Amino-Modifier is also available from Clontech Laboratories Inc. (Palo Alto, CA). In addition, a nucleotide analog bearing a linking group pre-attached to the nucleoside

15 is commercially available from Glen Research Corporation under the tradename "Amino-Modifier-dT." This nucleoside-linking group reagent, a uridine derivative having an [N(7-trifluoroacetyl aminoheptyl)3-acrylamido] substituent group at the 5 position of the pyrimidine ring, is synthesized as per

20 the procedure of Jablonski, E., Moomaw, E.w., Tullis, R.H. and Ruth, J.L. (1986) *Nucleic Acid Research*, 14:6115. The present invention also includes as nucleoside analogs adenine nucleosides functionalized to include a linker on the N6 purine amino group, guanine nucleosides functionalized to

25 include a linker at the exocyclic N2 purine amino group, and cytosine nucleosides functionalized to include a linker on either the N4 pyrimidine amino group or the 5 pyrimidine position.

The oligonucleotide-folate conjugates of the present

30 invention also include conjugates of oligonucleotides and linked nucleosides wherein the folate is attached directly onto the nucleoside or nucleotide without the intermediacy of a linker group. This attachment of folate may be performed at either one or more of the 2'-, 3'-, 5'-, nucleobase or

35 internucleotide linkage positions of the oligonucleotide or

linked nucleosides of the invention. In the compounds and methods of the invention described herein, such structures are represented by those wherein variable  $W_1$  is O, NH or S. The attachment of the folate may be performed at either of  
5 the  $\alpha$ -carboxylate,  $\gamma$ -carboxylate, amino or oxo groups of the folate. Typical linkages may be as ester, amide, or carbamate groups.

The present invention employs oligonucleotides for use in antisense modulation of the function of DNA or messenger  
10 RNA (mRNA) encoding a protein the modulation of which is desired, and ultimately to regulate the amount of such a protein. Hybridization of an antisense oligonucleotide with its mRNA target interferes with the normal role of mRNA and causes a modulation of its function in cells. The functions  
15 of mRNA to be interfered with include all vital functions such as translocation of the RNA to the site for protein translation, actual translation of protein from the RNA, splicing of the RNA to yield one or more mRNA species, turnover or degradation of the mRNA and possibly even  
20 independent catalytic activity which may be engaged in by the RNA. The overall effect of such interference with mRNA function is modulation of the expression of a protein, wherein "modulation" means either an increase (stimulation) or a decrease (inhibition) in the expression of the protein.  
25 In the context of the present invention, inhibition is the preferred form of modulation of gene expression.

In the context of this invention, the term "oligonucleotide" refers to an oligomer or polymer of ribonucleic acid or deoxyribonucleic acid. This term  
30 includes oligonucleotides composed of naturally-occurring nucleobases, sugars and covalent intersugar (backbone) linkages as well as modified oligonucleotides having non-naturally-occurring portions which function similarly. Such modified or substituted oligonucleotides are often preferred  
35 over native forms because of desirable properties such as,

for example, enhanced cellular uptake, enhanced binding to target and increased stability in the presence of nucleases.

An oligonucleotide is a polymer of repeating units generically known as a nucleotides. An unmodified (naturally occurring) nucleotide has three components: (1) a nitrogenous base linked by one of its nitrogen atoms to (2) a 5-carbon cyclic sugar and (3) a phosphate, esterified to carbon 5 of the sugar. When incorporated into an oligonucleotide chain, the phosphate of a first nucleotide is also esterified to carbon 3 of the sugar of a second, adjacent nucleotide. The "backbone" of an unmodified oligonucleotide consists of (2) and (3), that is, sugars linked together by phosphodiester linkages between the carbon 5 (5') position of the sugar of a first nucleotide and the carbon 3 (3') position of a second, adjacent nucleotide. a "nucleoside" is the combination of (1) a nucleobase and (2) a sugar in the absence of (3) a phosphate moiety (Kornberg, a., *DNA Replication*, W.H. Freeman & Co., San Francisco, 1980, pages 4-7). The backbone of an oligonucleotide positions a series of bases in a specific order; the written representation of this series of bases, which is conventionally written in 5' to 3' order, is known as a nucleotide sequence.

Oligonucleotides may comprise nucleotide sequences sufficient in identity and number to effect specific hybridization with a particular nucleic acid. Such oligonucleotides which specifically hybridize to a portion of the sense strand of a gene are commonly described as "antisense." In the context of the invention, "hybridization" means hydrogen bonding, which may be Watson-Crick, Hoogsteen or reversed Hoogsteen hydrogen bonding, between complementary nucleotides. For example, adenine and thymine are complementary nucleobases which pair through the formation of hydrogen bonds. "Complementary," as used herein, refers to the capacity for precise pairing between two nucleotides. For example, if a nucleotide at a certain



position of an oligonucleotide is capable of hydrogen bonding with a nucleotide at the same position of a DNA or RNA molecule, then the oligonucleotide and the DNA or RNA are considered to be complementary to each other at that position. The oligonucleotide and the DNA or RNA are complementary to each other when a sufficient number of corresponding positions in each molecule are occupied by nucleotides which can hydrogen bond with each other. Thus, "specifically hybridizable" and "complementary" are terms which are used to indicate a sufficient degree of complementarity or precise pairing such that stable and specific binding occurs between the oligonucleotide and the DNA or RNA target. It is understood in the art that an oligonucleotide need not be 100% complementary to its target DNA sequence to be specifically hybridizable. An oligonucleotide is specifically hybridizable when binding of the oligonucleotide to the target DNA or RNA molecule interferes with the normal function of the target DNA or RNA to cause a decrease or loss of function, and there is a sufficient degree of complementarity to avoid non-specific binding of the oligonucleotide to non-target sequences under conditions in which specific binding is desired, i.e., under physiological conditions in the case of *in vivo* assays or therapeutic treatment, or in the case of *in vitro* assays, under conditions in which the assays are performed.

Antisense oligonucleotides are commonly used as research reagents, diagnostic aids, and therapeutic agents. For example, antisense oligonucleotides, which are able to inhibit gene expression with exquisite specificity, are often used by those of ordinary skill to elucidate the function of particular genes, for example to distinguish between the functions of various members of a biological pathway. This specific inhibitory effect has, therefore, been harnessed by those skilled in the art for research uses. Antisense oligonucleotides have also been used as diagnostic aids based

on their specific binding or hybridization to DNA or mRNA that are present in certain disease states and due to the high degree of sensitivity that hybridization based assays and amplified assays that utilize some of polymerase chain reaction afford. The specificity and sensitivity of oligonucleotides is also harnessed by those of skill in the art for therapeutic uses. For example, the following U.S. patents demonstrate palliative, therapeutic and other methods utilizing antisense oligonucleotides. U. S. Patent No. 5,135,917 provides antisense oligonucleotides that inhibit human interleukin-1 receptor expression. U.S. Patent No. 5,098,890 is directed to antisense oligonucleotides complementary to the *c-myc* oncogene and antisense oligonucleotide therapies for certain cancerous conditions. U.S. Patent No. 5,087,617 provides methods for treating cancer patients with antisense oligonucleotides. U.S. Patent No. 5,166,195 provides oligonucleotide inhibitors of Human Immunodeficiency Virus (HIV). U.S. Patent No. 5,004,810 provides oligomers capable of hybridizing to herpes simplex virus Vmw65 mRNA and inhibiting replication. U.S. Patent No. 5,194,428 provides antisense oligonucleotides having antiviral activity against influenza virus. U.S. Patent No. 4,806,463 provides antisense oligonucleotides and methods using them to inhibit HTLV-III replication. U.S. Patent No. 5,286,717 provides oligonucleotides having a complementary base sequence to a portion of an oncogene. U.S. Patent No. 5,276,019 and U.S. Patent No. 5,264,423 are directed to phosphorothioate oligonucleotide analogs used to prevent replication of foreign nucleic acids in cells. U.S. Patent No. 4,689,320 is directed to antisense oligonucleotides as antiviral agents specific to cytomegalovirus (CMV). U.S. Patent No. 5,098,890 provides oligonucleotides complementary to at least a portion of the mRNA transcript of the human *c-myc* gene. U.S. Patent No. 5,242,906 provides antisense oligonucleotides useful in the treatment of latent Epstein-

Barr virus (EBV) infections. Other examples of antisense oligonucleotides are provided herein.

The oligonucleotides used in the compositions of the present invention preferably comprise from about 8 to about 5 30 nucleotides. It is more preferred that such oligonucleotides comprise from about 15 to 25 nucleotides. As is known in the art, a nucleotide is a base-sugar combination suitably bound to an adjacent nucleotide through a phosphodiester, phosphorothioate or other covalent linkage. 10 In the context of this invention, the term "oligonucleotide" includes oligonucleotides composed of naturally-occurring nucleobases, sugars and covalent intersugar (backbone) linkages as well as oligonucleotides having non-naturally-occurring portions which function similarly. Such modified 15 or substituted oligonucleotides may be preferred over native forms because of desirable properties such as, for example, enhanced cellular uptake, enhanced binding to target and increased stability in the presence of nucleases.

Specific examples of some preferred modified 20 oligonucleotides envisioned for use in the oligonucleotide-folate conjugates of the present invention include oligonucleotides containing modified backbones or non-natural intersugar linkages. As defined in this specification, oligonucleotides having modified backbones include those that 25 retain a phosphorus atom in the backbone and those that do not have a phosphorus atom in the backbone. For the purposes of this specification, and as sometimes referenced in the art, modified oligonucleotides that do not have a phosphorus atom in their intersugar backbone can also be considered to 30 be oligonucleosides.

Specific oligonucleotide chemical modifications are described below. It is not necessary for all positions in a given compound to be uniformly modified, and in fact more than one of the following modifications may be incorporated 35 in a single antisense compound or even in a single residue

thereof, for example, at a single nucleoside within an oligonucleotide.

Preferred modified oligonucleotide backbones include, for example, phosphorothioates, chiral phosphorothioates, 5 phosphorodithioates, phosphotriesters, aminoalkylphosphotriesters, methyl and other alkyl phosphonates including 3'-alkylene phosphonates and chiral phosphonates, phosphinates, phosphoramidates including 3'-amino phosphoramidate and aminoalkylphosphoramidates, thionophosphoramidates, thiono- 10 alkylphosphonates, thionoalkylphosphotriesters, and boranophosphates having normal 3'-5' linkages, 2'-5' linked analogs of these, and those having inverted polarity wherein the adjacent pairs of nucleoside units are linked 3'-5' to 5'-3' or 2'-5' to 5'-2'. Various salts, mixed salts and free acid 15 forms are also included.

Representative United States Patents that teach the preparation of the above phosphorus atom containing linkages include, but are not limited to, U.S. Patents Nos. 3,687,808; 4,469,863; 4,476,301; 5,023,243; 5,177,196; 5,188,897; 20 5,264,423; 5,276,019; 5,278,302; 5,286,717; 5,321,131; 5,399,676; 5,405,939; 5,453,496; 5,455,233; 5,466,677; 5,476,925; 5,519,126; 5,536,821; 5,541,306; 5,550,111; 5,563,253; 5,571,799; 5,587,361; 5,625,050; and 5,697,248, certain of which are commonly owned with this application, 25 and each of which is herein incorporated by reference.

Preferred modified oligonucleotide backbones that do not include a phosphorus atom therein (i.e., oligonucleosides) have backbones that are formed by short chain alkyl or cycloalkyl intersugar linkages, mixed 30 heteroatom and alkyl or cycloalkyl intersugar linkages, or one or more short chain heteroatomic or heterocyclic intersugar linkages. These include those having morpholino linkages (formed in part from the sugar portion of a nucleoside); siloxane backbones; sulfide, sulfoxide and 35 sulfone backbones; formacetyl and thioformacetyl backbones;

methylene formacetyl and thioformacetyl backbones; alkene containing backbones; sulfamate backbones; methyleneimino and methylenehydrazino backbones; sulfonate and sulfonamide backbones; amide backbones; and others having mixed N, O, S and CH<sub>2</sub> component parts.

Representative United States patents that teach the preparation of the above oligonucleosides include, but are not limited to, U.S. Patents Nos. 5,034,506; 5,166,315; 5,185,444; 5,214,134; 5,216,141; 5,235,033; 5,264,562; 10 5,264,564; 5,405,938; 5,434,257; 5,466,677; 5,470,967; 5,489,677; 5,541,307; 5,561,225; 5,596,086; 5,602,240; 5,610,289; 5,602,240; 5,608,046; 5,610,289; 5,618,704; 5,623,070; 5,663,312; 5,633,360; 5,677,437; and 5,677,439, certain of which are commonly owned with this application, 15 and each of which is herein incorporated by reference.

In other preferred oligonucleotide mimetics, both the sugar and the intersugar linkage, i.e., the backbone, of the nucleotide units are replaced with novel groups. The base units are maintained for hybridization with an appropriate 20 nucleic acid target compound. One such oligomeric compound, an oligonucleotide mimetic that has been shown to have excellent hybridization properties, is referred to as a peptide nucleic acid (PNA). In PNA compounds, the sugar-backbone of an oligonucleotide is replaced with an amide 25 containing backbone, in particular an aminoethylglycine backbone. The nucleobases are retained and are bound directly or indirectly to aza nitrogen atoms of the amide portion of the backbone. Representative United States patents that teach the preparation of PNA compounds include, 30 but are not limited to, U.S. Patents Nos. 5,539,082; 5,714,331; and 5,719,262, each of which is herein incorporated by reference. Further teaching of PNA compounds can be found in Nielsen et al., *Science*, 1991, 254, 1497.

Some preferred embodiments of the present invention 35 may employ oligonucleotides with phosphorothioate backbones

and oligonucleosides with heteroatom backbones, and in particular  $-\text{CH}_2-\text{NH}-\text{O}-\text{CH}_2-$ ,  $-\text{CH}_2-\text{N}(\text{CH}_3)-\text{O}-\text{CH}_2-$  [known as a methylene (methylimino) or MMI backbone],  $-\text{CH}_2-\text{O}-\text{N}(\text{CH}_3)-\text{CH}_2-$ ,  $-\text{CH}_2-\text{N}(\text{CH}_3)-\text{N}(\text{CH}_3)-\text{CH}_2-$  and  $-\text{O}-\text{N}(\text{CH}_3)-\text{CH}_2-\text{CH}_2-$  [wherein the native phosphodiester backbone is represented as  $-\text{O}-\text{P}-\text{O}-\text{CH}_2-$ ] of the above referenced U.S. Patent 5,489,677, and the amide backbones of the above referenced U.S. Patent No. 5,602,240. Also preferred are oligonucleotides having morpholino backbone structures of the above-referenced U.S. Patent No. 5,034,506.

The oligonucleotides employed in the oligonucleotide-folate conjugates of the present invention may additionally or alternatively comprise nucleobase (often referred to in the art simply as "base") modifications or substitutions. As used herein, "unmodified" or "natural" nucleobases include the purine bases adenine (a) and guanine (G), and the pyrimidine bases thymine (T), cytosine © and uracil (U). Modified nucleobases include other synthetic and natural nucleobases such as 5-methylcytosine (5-me-C), 5-hydroxymethyl cytosine, xanthine, hypoxanthine, 2-aminoadenine, 6-methyl and other alkyl derivatives of adenine and guanine, 2-propyl and other alkyl derivatives of adenine and guanine, 2-thiouracil, 2-thiothymine and 2-thiocytosine, 5-halouracil and cytosine, 5-propynyl uracil and cytosine, 6-azo uracil, cytosine and thymine, 5-uracil (pseudouracil), 4-thiouracil, 8-halo, 8-amino, 8-thiol, 8-thioalkyl, 8-hydroxyl and other 8-substituted adenines and guanines, 5-halo particularly 5-bromo, 5-trifluoromethyl and other 5-substituted uracils and cytosines, 7-methylguanine and 7-methyladenine, 8-azaguanine and 8-azaadenine, 7-deazaguanine and 7-deazaadenine and 3-deazaguanine and 3-deazaadenine. Further nucleobases include those disclosed in United States Patent No. 3,687,808, those disclosed in the *Concise Encyclopedia Of Polymer Science And Engineering*, pages 858-859, Kroschwitz, J.I., ed. John Wiley & Sons, 1990, those

disclosed by Englisch et al., *Angewandte Chemie, International Edition*, 1991, 30, 613, and those disclosed by Sanghvi, Y.S., Chapter 15, *Antisense Research and Applications*, pages 289-302, Crooke, S.T. and Lebleu, B., ed., CRC Press, 1993. Certain of these nucleobases are particularly useful for increasing the binding affinity of the oligomeric compounds of the invention. These include 5-substituted pyrimidines, 6-azapyrimidines and N-2, N-6 and O-6 substituted purines, including 2-aminopropyladenine, 5-propynyluracil and 5-propynylcytosine. 5-methylcytosine substitutions have been shown to increase nucleic acid duplex stability by 0.6-1.2°C (*Id.*, pages 276-278) and are presently preferred base substitutions, even more particularly when combined with 2'-methoxyethyl sugar modifications.

Representative United States patents that teach the preparation of certain of the above noted modified nucleobases as well as other modified nucleobases include, but are not limited to, the above noted U.S. Patent 3,687,808, as well as U.S. Patents 4,845,205; 5,130,302; 5,134,066; 5,175,273; 5,367,066; 5,432,272; 5,457,187; 5,459,255; 5,484,908; 5,502,177; 5,525,711; 5,552,540; 5,587,469; 5,594,121; 5,596,091; 5,614,617; and 5,681,941, certain of which are commonly owned, and each of which is herein incorporated by reference, and commonly owned United States patent application 08/762,488, filed on December 10, 1996, also herein incorporated by reference.

The oligonucleotides employed in the oligonucleotide-folate conjugates of the present invention may additionally or alternatively comprise one or more substituted sugar moieties. Preferred oligonucleotides comprise one of the following at the 2' position: OH; F; O-, S-, or N-alkyl, O-, S-, or N-alkenyl, or O, S- or N-alkynyl, wherein the alkyl, alkenyl and alkynyl may be substituted or unsubstituted C<sub>1</sub> to C<sub>10</sub> alkyl or C<sub>2</sub> to C<sub>10</sub> alkenyl and alkynyl. Particularly preferred are O[(CH<sub>2</sub>)<sub>n</sub>O]<sub>m</sub>CH<sub>3</sub>, O(CH<sub>2</sub>)<sub>n</sub>OCH<sub>3</sub>, O(CH<sub>2</sub>)<sub>n</sub>NH<sub>2</sub>, O(CH<sub>2</sub>)<sub>n</sub>CH<sub>3</sub>,

O(CH<sub>2</sub>)<sub>n</sub>ONH<sub>2</sub>, and O(CH<sub>2</sub>)<sub>n</sub>ON[(CH<sub>2</sub>)<sub>n</sub>CH<sub>3</sub>]<sub>2</sub>, where n and m are from 1 to about 10. Other preferred oligonucleotides comprise one of the following at the 2' position: C<sub>1</sub> to C<sub>10</sub> lower alkyl, substituted lower alkyl, alkaryl, aralkyl, O-alkaryl or O-aralkyl, SH, SCH<sub>3</sub>, OCN, Cl, Br, CN, CF<sub>3</sub>, OCF<sub>3</sub>, SOCH<sub>3</sub>, SO<sub>2</sub>CH<sub>3</sub>, ONO<sub>2</sub>, NO<sub>2</sub>, N<sub>3</sub>, NH<sub>2</sub>, heterocycloalkyl, heterocycloalkaryl, aminoalkylamino, polyalkylamino, substituted silyl, an RNA cleaving group, a reporter group, an intercalator, a group for improving the pharmacokinetic properties of an oligonucleotide, or a group for improving the pharmacodynamic properties of an oligonucleotide, and other substituents having similar properties. a preferred modification includes 2'-methoxyethoxy [2'-O-CH<sub>2</sub>CH<sub>2</sub>OCH<sub>3</sub>, also known as 2'-O-(2-methoxyethyl) or 2'-MOE] (Martin et al., *Helv. Chim. Acta*; 1995, 78, 486), i.e., an alkoxyalkoxy group. a further preferred modification includes 2'-dimethylaminooxyethoxy, i.e., a O(CH<sub>2</sub>)<sub>2</sub>ON(CH<sub>3</sub>)<sub>2</sub> group, also known as 2'-DMAOE, as described in co-owned United States patent application Serial Number 09/016,520, filed on January 30, 1998, the contents of which are herein incorporated by reference.

Other preferred modifications include 2'-methoxy (2'-O-CH<sub>3</sub>), 2'-aminopropoxy (2'-OCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub>) and 2'-fluoro (2'-F). Similar modifications may also be made at other positions on the oligonucleotide, particularly the 3' position of the sugar on the 3' terminal nucleotide or in 2'-5' linked oligonucleotides and the 5' position of 5' terminal nucleotide. Oligonucleotides may also have sugar mimetics such as cyclobutyl moieties in place of the pentofuranosyl sugar. Representative United States patents that teach the preparation of such modified sugars structures include, but are not limited to, U.S. Patents Nos. 4,981,957; 5,118,800; 5,319,080; 5,359,044; 5,393,878; 5,446,137; 5,466,786; 5,514,785; 5,519,134; 5,567,811; 5,576,427; 5,591,722; 5,597,909; 5,610,300; 5,627,053; 5,639,873; 5,646,265; 5,658,873; 5,670,633; and 5,700,920, certain of which are



commonly owned, and each of which is herein incorporated by reference, and commonly owned United States patent application 08/468,037, filed on June 5, 1995, also herein incorporated by reference.

5 Additional modifications may also be made at other positions on the oligonucleotide, particularly the 3' position of the sugar on the 3' terminal nucleotide and the 5' position of 5' terminal nucleotide. For example, one additional modification of the oligonucleotide-folate  
10 conjugates of the present invention involves chemically linking to the oligonucleotide one or more additional non-folate moieties or conjugates which enhance the activity, cellular distribution or cellular uptake of the oligonucleotide. Such moieties include but are not limited  
15 to lipid moieties such as a cholesterol moiety (Letsinger et al., *Proc. Natl. Acad. Sci. USA*, 1989, 86, 6553), cholic acid (Manoharan et al., *Bioorg. Med. Chem. Lett.*, 1994, 4, 1053), a thioether, e.g., hexyl-S-tritylthiol (Manoharan et al., *Ann. N.Y. Acad. Sci.*, 1992, 660, 306; Manoharan et al.,  
20 *Bioorg. Med. Chem. Lett.*, 1993, 3, 2765), a thiocholesterol (Oberhauser et al., *Nucl. Acids Res.*, 1992, 20, 533), an aliphatic chain, e.g., dodecandiol or undecyl residues (Saison-Behmoaras et al., *EMBO J.*, 1991, 10, 111; Kabanov et al., *FEBS Lett.*, 1990, 259, 327; Svinarchuk et al.,  
25 *Biochimie*, 1993, 75, 49), a phospholipid, e.g., di-hexadecyl-rac-glycerol or triethylammonium 1,2-di-O-hexadecyl-rac-glycero-3-H-phosphonate (Manoharan et al., *Tetrahedron Lett.*, 1995, 36, 3651; Shea et al., *Nucl. Acids Res.*, 1990, 18, 3777), a polyamine or a polyethylene glycol chain (Manoharan  
30 et al., *Nucleosides & Nucleotides*, 1995, 14, 969), or adamantane acetic acid (Manoharan et al., *Tetrahedron Lett.*, 1995, 36, 3651), a palmityl moiety (Mishra et al., *Biochim. Biophys. Acta*, 1995, 1264, 229), or an octadecylamine or hexylamino-carbonyl-oxycholesterol moiety (Crooke et al., *J.*

*Pharmacol. Exp. Ther.*, 1996, 277, 923).

Representative United States patents that teach the preparation of such oligonucleotide conjugates include, but are not limited to, U.S. Patents Nos. 4,828,979; 4,948,882; 5 5,218,105; 5,525,465; 5,541,313; 5,545,730; 5,552,538; 5,578,717; 5,580,731; 5,580,731; 5,591,584; 5,109,124; 5,118,802; 5,138,045; 5,414,077; 5,486,603; 5,512,439; 5,578,718; 5,608,046; 4,587,044; 4,605,735; 4,667,025; 4,762,779; 4,789,737; 4,824,941; 4,835,263; 4,876,335; 10 4,904,582; 4,958,013; 5,082,830; 5,112,963; 5,214,136; 5,082,830; 5,112,963; 5,214,136; 5,245,022; 5,254,469; 5,258,506; 5,262,536; 5,272,250; 5,292,873; 5,317,098; 5,371,241; 5,391,723; 5,416,203; 5,451,463; 5,510,475; 5,512,667; 5,514,785; 5,565,552; 5,567,810; 5,574,142; 15 5,585,481; 5,587,371; 5,595,726; 5,597,696; 5,599,923; 5,599,928 and 5,688,941, certain of which are commonly owned, and each of which is herein incorporated by reference.

The present invention also includes compositions employing antisense compounds which are chimeric compounds. 20 "Chimeric" antisense compounds or "chimeras," in the context of this invention, are antisense compounds, particularly oligonucleotides, which contain two or more chemically distinct regions, each made up of at least one monomer unit, i.e., a nucleotide in the case of an oligonucleotide 25 compound. These oligonucleotides typically contain at least one region wherein the oligonucleotide is modified so as to confer upon the oligonucleotide increased resistance to nuclease degradation, increased cellular uptake, and/or increased binding affinity for the target nucleic acid. An 30 additional region of the oligonucleotide may serve as a substrate for enzymes capable of cleaving RNA:DNA or RNA:RNA hybrids. By way of example, RNase H is a cellular endonuclease which cleaves the RNA strand of an RNA:DNA duplex. Activation of RNase H, therefore, results in 35 cleavage of the RNA target, thereby greatly enhancing the

efficiency of oligonucleotide inhibition of gene expression. Consequently, comparable results can often be obtained with shorter oligonucleotides when chimeric oligonucleotides are used, compared to phosphorothioate oligodeoxynucleotides hybridizing to the same target region. Cleavage of the RNA target can be routinely detected by gel electrophoresis and, if necessary, associated nucleic acid hybridization techniques known in the art. RNase H-mediated target cleavage is distinct from the use of ribozymes to cleave nucleic acids, and ribozymes are not comprehended by the present invention.

By way of example, such "chimeras" may be "gapmers," i.e., oligonucleotides in which a central portion (the "gap") of the oligonucleotide serves as a substrate for, e.g., RNase H, and the 5' and 3' portions (the "wings") are modified in such a fashion so as to have greater affinity for, or stability when duplexed with, the target RNA molecule but are unable to support nuclease activity (e.g., 2'-fluoro- or 2'-methoxyethoxy- substituted). Other chimeras include "hemimers," that is, oligonucleotides in which the 5' portion of the oligonucleotide serves as a substrate for, e.g., RNase H, whereas the 3' portion is modified in such a fashion so as to have greater affinity for, or stability when duplexed with, the target RNA molecule but is unable to support nuclease activity (e.g., 2'-fluoro- or 2'-methoxyethoxy-substituted), or vice-versa.

a number of chemical modifications to oligonucleotides that confer greater oligonucleotide:RNA duplex stability have been described by Freier et al. (*Nucl. Acids Res.*, 1997, 25, 4429). Such modifications are preferred for the RNase H-refractory portions of chimeric oligonucleotides and may generally be used to enhance the affinity of an antisense compound for a target RNA.

Chimeric antisense compounds of the invention may be formed as composite structures of two or more

oligonucleotides, modified oligonucleotides, oligonucleosides and/or oligonucleotide mimetics as described above. Such compounds have also been referred to in the art as hybrids or gapmers. Representative United States patents that teach the preparation of such hybrid structures include, but are not limited to, U.S. Patents Nos. 5,013,830; 5,149,797; 5,220,007; 5,256,775; 5,366,878; 5,403,711; 5,491,133; 5,565,350; 5,623,065; 5,652,355; 5,652,356; and 5,700,922, certain of which are commonly owned, and each of which is herein incorporated by reference, and commonly owned and allowed United States patent application serial number 08/465,880, filed on June 6, 1995, also herein incorporated by reference.

The present invention also includes compositions employing oligonucleotides that are substantially chirally pure with regard to particular positions within the oligonucleotides. Examples of substantially chirally pure oligonucleotides include, but are not limited to, those having phosphorothioate linkages that are at least 75% Sp or Rp (Cook et al., U.S. Patent No. 5,587,361) and those having substantially chirally pure (Sp or Rp) alkylphosphonate, phosphoramidate or phosphotriester linkages (Cook, U.S. Patents Nos. 5,212,295 and 5,521,302).

The present invention further encompasses oligonucleotide-folate conjugates employing ribozymes. Synthetic RNA molecules and derivatives thereof that catalyze highly specific endoribonuclease activities are known as ribozymes. (See, generally, U.S. Patent No. 5,543,508 to Haseloff et al., issued August 6, 1996, and U.S. Patent No. 5,545,729 to Goodchild et al., issued August 13, 1996.) The cleavage reactions are catalyzed by the RNA molecules themselves. In naturally occurring RNA molecules, the sites of self-catalyzed cleavage are located within highly conserved regions of RNA secondary structure (Buzayan et al., *Proc. Natl. Acad. Sci. U.S.A.*, 1986, 83, 8859; Forster et

al., Cell, 1987, 50, 9). Naturally occurring autocatalytic RNA molecules have been modified to generate ribozymes which can be targeted to a particular cellular or pathogenic RNA molecule with a high degree of specificity. Thus, ribozymes  
5 serve the same general purpose as antisense oligonucleotides (i.e., modulation of expression of a specific gene) and, like oligonucleotides, are nucleic acids possessing significant portions of single-strandedness. That is, ribozymes have substantial chemical and functional identity with  
10 oligonucleotides and are thus considered to be equivalents for purposes of the present invention.

The oligonucleotides used in the conjugates of the present invention may be conveniently and routinely made through the well-known technique of solid phase synthesis.  
15 Equipment for such synthesis is sold by several vendors including, for example, Applied Biosystems (Foster City, CA). Any other means for such synthesis known in the art may additionally or alternatively be employed. It is also known to use similar techniques to prepare other oligonucleotides  
20 such as the phosphorothioates and alkylated derivatives.

Teachings regarding the synthesis of particular modified oligonucleotides may be found in the following U.S. patents or pending patent applications, each of which is commonly assigned with this application: U.S. Patents Nos.  
25 5,138,045 and 5,218,105, drawn to polyamine conjugated oligonucleotides; U.S. Patent No. 5,212,295, drawn to monomers for the preparation of oligonucleotides having chiral phosphorus linkages; U.S. Patents Nos. 5,378,825 and 5,541,307, drawn to oligonucleotides having modified  
30 backbones; U.S. Patent No. 5,386,023, drawn to backbone modified oligonucleotides and the preparation thereof through reductive coupling; U.S. Patent No. 5,457,191, drawn to modified nucleobases based on the 3-deazapurine ring system and methods of synthesis thereof; U.S. Patent No. 5,459,255,  
35 drawn to modified nucleobases based on N-2 substituted

purines; U.S. Patent No. 5,521,302, drawn to processes for preparing oligonucleotides having chiral phosphorus linkages; U.S. Patent No. 5,539,082, drawn to peptide nucleic acids; U.S. Patent No. 5,554,746, drawn to oligonucleotides having  
5  $\beta$ -lactam backbones; U.S. Patent No. 5,571,902, drawn to methods and materials for the synthesis of oligonucleotides; U.S. Patent No. 5,578,718, drawn to nucleosides having alkylthio groups, wherein such groups may be used as linkers to other moieties attached at any of a variety of positions  
10 of the nucleoside; U.S. Patents Nos. 5,587,361 and 5,599,797, drawn to oligonucleotides having phosphorothioate linkages of high chiral purity; U.S. Patent No. 5,506,351, drawn to processes for the preparation of 2'-O-alkyl guanosine and related compounds, including 2,6-diaminopurine compounds;  
15 U.S. Patent No. 5,587,469, drawn to oligonucleotides having N-2 substituted purines; U.S. Patent No. 5,587,470, drawn to oligonucleotides having 3-deazapurines; U.S. Patents Nos. 5,223,168, issued June 29, 1993, and 5,608,046, both drawn to conjugated 4'-desmethyl nucleoside analogs; U.S. Patent Nos.  
20 5,602,240, and 5,610,289, drawn to backbone modified oligonucleotide analogs; and U.S. patent application Serial No. 08/383,666, filed February 3, 1995, and U.S. Patent No. 5,459,255, drawn to, *inter alia*, methods of synthesizing 2'-fluoro-oligonucleotides.

25 In the oligonucleotide-folate conjugates and folate molecule-bearing sequence-specific linked nucleosides of the present invention, the nucleosides are assembled on a suitable DNA synthesizer utilizing standard nucleotide precursors, or nucleotide conjugate precursors that already  
30 bear the linking moiety, or folate-nucleoside conjugate precursors that already bear the folate molecule.

When using nucleotide conjugate precursors that already bear a linking moiety, the synthesis of the sequence-specific linked nucleosides is typically completed, and the  
35 folate molecule is then reacted with the linking moiety to

form the oligonucleotide-folate conjugate. This approach to the synthesis of oligonucleotide conjugates bearing a variety of molecules such as steroids, vitamins, lipids and reporter molecules has previously been described. See Manoharan et al., PCT Application WO 93/07883. In a preferred embodiment the oligonucleotides or linked nucleosides of the present invention are synthesized by an automated synthesizer using phosphoramidites derived from folate-nucleoside conjugates in addition to the standard phosphoramidites used in oligonucleotide synthesis.

In application serial No. US91/00243, application serial No. 463,358 and application serial No. 566,977, all incorporated herein by reference, it is reported that incorporation of a 2'-O-methyl, 2'-O-ethyl, 2'-O-propyl, 2'-O-allyl, 2'-O-aminoalkyl or 2'-deoxy-2'-fluoro group in nucleosides of an oligonucleotide confers enhanced hybridization properties to the oligonucleotide. It is further reported that oligonucleotides containing phosphorothioate backbones have enhanced nuclease stability. Thus, functionalized, linked nucleosides of the invention can be augmented to further include either or both a phosphorothioate backbone or a 2'-O-methyl, 2'-O-ethyl, 2'-O-propyl, 2'-O-aminoalkyl, 2'-O-allyl or 2'-deoxy-2'-fluoro group thereon.

In some preferred embodiments, functionalized nucleoside sequences of the invention possessing an amino group at the 5'-terminus is prepared using a DNA synthesizer, and then reacted with an active ester derivative of folic acid. Active ester derivatives are well known to those skilled in the art. Representative active esters include N-hydrosuccinimide esters, tetrafluorophenolic esters, pentafluorophenolic esters and pentachlorophenolic esters. The reaction of the amino group and the active ester produces an oligonucleotide in which folic acid is attached to the 5'-position through a linking group. The amino group at the 5'-

terminus can conveniently be prepared utilizing the above noted 5'-Amino-Modifier C6 reagent. In a preferred embodiment, folate molecules may be conjugated to oligonucleotides at the 5'-position by the use of a folate-nucleoside phosphoramidite wherein the folate is linked to the 5'-hydroxy group directly or indirectly via a linker. Such folate-nucleoside phosphoramidites are typically used at the end of an automated synthesis procedure to provide an oligonucleotide-folate conjugate bearing the folate molecule at the 5'-terminus.

In further preferred embodiments, functionalized nucleoside sequences of the invention can be prepared wherein folic acid is attached to the 3'-terminal amino group using a 3'-amino modified controlled pore glass (sold by Clontech Laboratories Inc., Palo Alto, CA) and subsequent attachment of the folic acid is achieved by reaction with a folic acid active ester.

In another preferred embodiment of the present invention, the folate may be attached to the oligonucleotide at the 3'-terminus through the use of an appropriate multifunctional linker such as a universal linker. In this case the folate is first derivatized with the universal linker and this conjugate then loaded onto a solid support. Subsequent synthesis of nucleic acids or oligonucleotides on this solid support affords upon cleavage and deprotection the oligonucleotide-folate conjugate bearing a folate molecule at the 3'-terminus.

In still further preferred embodiments, functionalized sequences of nucleosides and oligonucleotide-folate conjugates of the present invention can be prepared wherein the folate molecule is attached either directly or via a linking group to any one of the atoms of the nucleobase of any of the nucleoside units of the oligonucleotide. Thus, one or more folate molecules may be attached to the nucleobase at the 3'-terminus, the 5'-terminus or any



position in between. Such attachment can be accomplished, for example, by chemistries described in the literature, and mentioned above. The preferred mode of attachment of folate molecules to nucleobases is via the intermediacy of an

5 appropriate linker present on a nucleoside precursor. The folate-nucleoside conjugate is then phosphitylated at the 3'-position to afford a folate-nucleoside conjugate phosphoramidite which may be used subsequently as a building block together with traditional nucleoside phosphoramidites  
10 for the automated synthesis of oligonucleotides. The number and location of insertion of such folate nucleotide conjugate phosphoramidites will then dictate the number and location of folate molecules present in the synthesized oligonucleotide-folate conjugate of the present invention.

15 The present invention also provides oligonucleotide-folate conjugates wherein the folate molecule is attached to one of the atoms of the internucleotide linkage. One typical internucleotide linkage in nucleic acids and oligonucleotides is the phosphodiester linkage. Numerous modified  
20 internucleotide linkages are known in the art including, but not limited to, phosphorothioate, methyl phosphonate, and phosphordithioate, as described above. Folate molecules may be conjugated at one of the atoms of such internucleotide linkages with or without the intermediacy of a linking group.  
25 Attachment of the folate molecule may be accomplished in accordance with the methods of the invention either during the preparation of the nucleoside building block such as the phosphoramidite or may be performed during the formation of the internucleotide linkage during oligonucleotide synthesis.

30 In addition to the multiple sites on an oligonucleotide where a folate molecule may be attached, there exist multiple sites on the folate molecule through which the conjugation may occur. In accordance with the present invention the folic acid component of the  
35 oligonucleotide-folic acid conjugate can be attached at one

of, but not limited to, the a-carboxylate, the g-carboxylate, or amino group of the folate. Most preferred among the various available sites for conjugation are the a-carboxylate and the g-carboxylate of the folic acid molecule, or  
5 derivative thereof. Conjugation is typically performed using an aminoalkyl linker via an amide bond.

In further preferred embodiments of the invention, the folate molecule is attached at multiple sites on one oligonucleotide. For example, oligonucleotide-folate  
10 conjugates can be prepared wherein one or more folates are attached to both ends of a linked nucleoside sequence. Preferably such a structure is prepared by reacting a 3',5'-diamino sequence with a folic acid active ester. The required oligonucleoside sequence can be synthesized, for  
15 example, utilizing the 3'-Amino-Modifier and the 5'-Amino-Modifier C6 (or Aminolink-2) reagents noted above or by utilizing the above noted 3'-amino modified controlled pore glass reagent in combination with the 5'-Amino-Modifier C2 (or Aminolink-2) reagents. Alternatively, such multiply  
20 conjugated oligonucleotides may readily be synthesized according to the methods of the invention using an appropriate folate-nucleoside conjugate phosphoramidites as and where needed in a given oligonucleotide sequence during automated synthesis.

25 In still further preferred embodiments of the invention, an oligonucleoside sequence bearing an aminolinker at the 2'-position of one or more selected nucleosides is prepared using a suitably functionalized nucleotide such as, for example, 5'-dimethoxytrityl-2'-O-(ε-phthalimidylamino-  
30 pentyl)-2'-deoxyadenosine-3'-N,N-diisopropyl-cyanoethoxy phosphoramidite. See the above referenced patent applications serial numbers US91/00243, 566,977 and 463,358. Preferably, the nucleotide or nucleotides are attached to the folic acid by reaction with an active ester or a thioisocyanate thereof,  
35 at one or more of the nucleoside components of the

oligonucleotide.

In yet further preferred embodiments, functionalized nucleoside sequences of the invention can be prepared wherein the heterocyclic base of one or more nucleosides can be  
5 linked to a folate molecule. for example, utilizing 5'-O-dimethoxytrityl-5-[N(7-trifluoroacetyl-aminoheptyl)-3-acrylamido]-2'-deoxyuridine 3'-O-(methyl N,N-diisopropyl)phosphor-  
amide as described by Jablonski et. al. *supra* (also commercially available from Glen Research) the desired  
10 nucleoside, functionalized to incorporate a linking group on its heterocyclic base, is incorporated into the linked nucleoside sequence using a DNA synthesizer.

In further functionalized linked nucleosides of the invention, conjugation (or linking) of folate molecules is  
15 achieved by conjugation of the folate to the above described amino linking group on the nucleoside. This can be effected in several ways. For example, a folate-nucleoside conjugate of the invention can be prepared by conjugation of the folate molecule to the nucleoside using EDC/sulfo-NHS (i.e. 1-ethyl-  
20 3(3-dimethylaminopropylcarbodiimide/N-hydroxysulfosuccinimide) to conjugate the carboxylate function of the folate with the amino function of the linking group on the - nucleoside.

Oligonucleotide-folate conjugates of the present  
25 invention may be prepared by conjugation of the folate molecule to the nucleoside sequence via a heterobifunctional linker such as *m*-maleimidobenzoyl-N-hydroxysulfosuccinimide ester (MBS) or succinimidyl 4-(N-maleimidomethyl)cyclohexane-1-carboxylate (SMCC), to link a nucleophilic position,  
30 preferably a thiol, on the folate molecule to the amino function of the linking group on nucleoside sequence. By this mechanism, an oligonucleoside-maleimide conjugate is formed by reaction of the amino group of the linker on the linked nucleosides with the MBS or SMCC maleimide linker.  
35 The conjugate is then reacted with folate molecules,

preferably those that possess a thiol functionality.

Alternatively, an oligonucleotide-folate conjugate can be prepared by conjugation of the folate molecule to the oligonucleotide or nucleoside via a homobifunctional linker such as disuccinimidyl suberate (DSS), to link an amino function on the folate to the amino group of a linker on the oligonucleotide sequence. By this mechanism, an oligonucleoside-succinimidyl conjugate is formed by reaction of the amino group of the linker on the nucleoside sequence with a disuccinimidyl suberate linker. The disuccinimidyl suberate linker couples with the amine linker on the nucleoside to extend the size of the linker. The extended linker is then reacted with an amino group of the folate molecule.

A number of non-folate molecules have been conjugated to oligonucleotides in order to enhance the activity, cellular distribution or cellular uptake of the oligonucleotide, and procedures for performing such conjugations are available in the scientific literature. Such non-folate moieties have included lipid moieties such as cholesterol (Letsinger et al., *Proc. Natl. Acad. Sci. USA*, 1989, 86, 6553), cholic acid (Manoharan et al., *Bioorg. Med. Chem. Lett.*, 1994, 4, 1053), a thioether, e.g., hexyl-S-tritylthiol (Manoharan et al., *Ann. N.Y. Acad. Sci.*, 1992, 660, 306; Manoharan et al., *Bioorg. Med. Chem. Lett.*, 1993, 3, 2765), a thiocholesterol (Oberhauser et al., *Nucl. Acids Res.*, 1992, 20, 533), an aliphatic chain, e.g., dodecandiol or undecyl residues (Saison-Behmoaras et al., *EMBO J.*, 1991, 10, 111; Kabanov et al., *FEBS Lett.*, 1990, 259, 327; Svinarchuk et al., *Biochimie*, 1993, 75, 49), a phospholipid, e.g., di-hexadecyl-rac-glycerol or triethylammonium 1,2-di-O-hexadecyl-rac-glycero-3-H-phosphonate (Manoharan et al., *Tetrahedron Lett.*, 1995, 36, 3651; Shea et al., *Nucl. Acids Res.*, 1990, 18, 3777), a polyamine or a polyethylene glycol

chain (Manoharan et al., *Nucleosides & Nucleotides*, 1995, 14, 969), or adamantane acetic acid (Manoharan et al., *Tetrahedron Lett.*, 1995, 36, 3651), a palmityl moiety (Mishra et al., *Biochim. Biophys. Acta*, 1995, 1264, 229), or an

5 octadecylamine or hexylamino-carbonyl-oxycholesterol moiety (Crooke et al., *J. Pharmacol. Exp. Ther.*, 1996, 277, 923). Representative United States patents that teach the preparation of such oligonucleotide conjugates have been listed above. Typical conjugation protocols involve the

10 synthesis of oligonucleotides bearing an aminolinker at one or more positions of the sequence. The amino group is then reacted with the molecule being conjugated using appropriate coupling or activating reagents. The conjugation reaction may be performed either with the oligonucleotide still bound

15 to the solid support or following cleavage of the oligonucleotide in solution phase. Purification of the oligonucleotide conjugate by HPLC typically affords the pure conjugate.

Alternatively, the molecule being conjugated may be

20 converted into a building block such as a phosphoramidite via an alcohol group present in the molecule or by attachment of a linker bearing an alcohol group that may be phosphitylated.

Each of these approaches may be used for the synthesis of oligonucleotide-folate conjugates. Aminolinked

25 oligonucleotides may be coupled directly with folic acid via the use of coupling reagents or following activation of the folic acid as an NHS or pentfluorophenolate ester. While this procedure does afford the oligonucleotide-folic acid conjugate, the yield is low and the reaction affords a

30 mixture of the a-conjugate and g-conjugate formed via attachment of the a-carboxylic acid and the g-carboxylic acid of folic acid to the amino group of the oligonucleotide. Folic acid phosphoramidites may be synthesized via the attachment of an aminohexanol linker to one of the carboxyl

35 groups followed by phosphitylation of the terminal alcohol

functionality. Other linkers such as cysteamine may also be utilized for conjugation to a chloroacetyl linker present on a synthesized oligonucleotide.

The foregoing approaches can afford oligonucleotide-  
5 folate conjugates. However, they are limited by the nature of the folate that may be coupled because of the common insolubility of many folates and folate analogs. Further, these approaches lack regioselectivity, because of the two similar carboxylic groups present in many folates, and will  
10 often afford a mixture of a- and g-conjugates that may result in variability in the biological effects observed when studying such conjugates. Because these conjugations are post-synthetic, that is, after automated synthesis of the oligonucleotide has been performed, these approaches are  
15 limited to providing conjugates at only those positions that may be derivatized during automated synthesis and subsequently reacted with folate.

The methods of the present invention also provide synthetic paths that circumvent the solubility and  
20 regioselectivity problems of such conventional conjugations of folic acid and related folates, by using a novel conjugation approach that constructs the folate from its three major components during the process of conjugation.

Most folates are comprised of three portions: a  
25 pterin, p-aminobenzoic acid and glutamic acid which together form the pteroyl-p-aminobenzoylglutamate structure of folic acid. In preferred embodiments, the methods of the present invention provide for the synthesis of oligonucleotide-folate conjugates via first attachment of appropriately protected  
30 glutamic acid or other amino acids such as homoglutamic acid and aspartic acid to a nucleoside precursor or oligonucleotide. This attachment is preferably performed via an acylation reaction at the amine functionality of an amino linker present on the nucleoside or oligonucleotide. An  
35 appropriate pteroyl-p-amino benzoic acid is then condensed

with the pendant amino group of the previously conjugated glutamic acid. This two step conjugation reaction is affords significant benefits over current synthetic methods, including the elimination of all the problems associated with poor solubility of folates during conjugation reactions. In addition, because of the availability of many different selectively protected glutamate and other amino acids, the methods of this invention provide tremendous regioselectivity and control over the nature of the oligonucleotide-folate conjugate that may be synthesized.

The methods of the invention provide for the preparation of a variety of folate compounds wherein the glutamic acid moiety of the folic acid is replaced with another amino acid, preferably a naturally occurring amino acid. As used herein, the term "naturally occurring amino acid" is intended to mean amino acids found in nature, for example the  $\alpha$ -amino acids having the formula  $\text{HOOC}-\text{CH}(\text{sidechain})-\text{NH}_2$  that occur in living organisms. Representative sidechains of naturally occurring amino acids include the glutamyl side chain  $(-\text{CH}_2)_2-\text{COOH}$ , the aspartyl sidechain  $(-\text{CH}_2-\text{COOH})$ . Further representative sidechains of naturally occurring amino acid can be found in Lehninger, *Biochemistry*, Second Edition, Worth Publishers, Inc, 1975, pages 73-75, which is incorporated by reference herein.

The methods of the present invention allow for the selective synthesis of each of the two regioisomeric oligonucleotide-folate conjugates that may be formed by reaction of folic acid at either the  $\alpha$ -carboxylate or the  $\gamma$ -carboxylate, by systematically constructing the folate conjugate. Because glutamic acid is available in a number of different selectively protected derivatives, the  $\alpha$ -carboxylate of glutamic acid may be selectively reacted with a pendant amino group on a nucleoside by using the  $\gamma$ -t-butyl ester of N-Fmoc glutamic acid. Likewise, the  $\gamma$ -carboxylate of glutamic acid may be selectively reacted with a pendant amino

group on a nucleoside by using the *a*-*t*-butyl ester of N-Fmoc glutamic acid. Subsequent reaction with pteronic acid, according to the methods of the invention affords selectively both regioisomers of the oligonucleotide-folate conjugate as  
5 discrete compounds of the methods of the invention.

In one preferred embodiment of the methods of the invention, the preparation of oligonucleotide-folate conjugates of the present invention commences with the selection of appropriate precursor molecules upon which to  
10 construct the folate molecule. Typically the precursor is an appropriately protected derivative of the commonly used nucleosides. For example, the synthetic precursors for the synthesis of the oligonucleotide-folate conjugates of the present invention include, but are not limited to, 2'-*w*-  
15 aminoalkoxy-5'-ODMT-nucleosides, 2'-*w*-aminoalkylamino-5'-ODMT-nucleosides, 5'-*w*-aminoalkoxy-2'-deoxy-nucleosides, 5'-*w*-aminoalkoxy-2'-protected-nucleosides, 3'-*w*-aminoalkoxy-5'-ODMT-nucleosides, and 3'-aminoalkylamino-5'-ODMT-nucleosides that may be further protected in the nucleobase portion of.  
20 the molecule. The use of such precursors is anticipated to afford oligonucleotide-folate conjugates where attachment is at one of many possible sites such as the 2', 3' or 5' position of one or more of the nucleoside components of oligonucleotides. Methods for the synthesis of such  
25 aminolinked protected nucleoside precursors are known to the art skilled and are available in the literature.

In accordance with preferred embodiments of the methods of the invention, the aminolinked protected nucleoside precursors are first reacted with an appropriately  
30 protected amino acid, preferably glutamic acid, more preferably the *a*-*t*-butyl ester of N-Fmoc glutamic acid or the *g*-*t*-butyl ester of N-Fmoc glutamic acid. Alternate protecting groups that may be used for the pendant carboxylate group include the *O*-allyl ester. Coupling of the  
35 free carboxylate of the glutamic acid with the available



amino group pendant from the nucleoside precursor is performed, for example, using dicyclohexylcarbodiimide (DCC) and HONB as activating reagent in anhydrous DMF. In some preferred embodiments, 2'-O-(hexylamino)-5'-ODMT-5-methyl-  
5 uridine is reacted with N-Fmoc-a-t-butyl-L-glutamic acid affords the corresponding nucleoside-g-glutamic acid conjugate. Likewise reaction of 2'-O-(hexylamino)-5'-ODMT-uridine with N-Fmoc-g-t-butyl-L-glutamic acid affords the corresponding nucleoside-a-glutamic acid conjugate. The  
10 methods of the present invention, in addition to their utility in the preparation of such nucleoside-amino acid conjugates, are also applicable for the synthesis of a variety of other conjugates derived from other nucleosides and amino acids.

15 In some preferred embodiments, nucleoside-glutamic acid conjugates prepared according to the methods of the present invention are next treated with a deprotecting reagent such as, but not limited to, a mixture of 0.1 M DBU in acetonitrile or a 10% solution of piperidine in  
20 acetonitrile, to cleave the pendant amine protecting group such as the Fmoc group (i.e., "deprotection"). In further preferred embodiments, the intermediate nucleoside-glutamic acid conjugates of the invention are then reacted with a pteroyl-p-aminobenzoate to provide the folate molecule.

25 The pteroyl-p-aminobenzoate intermediate necessary for the construction of folate conjugates may also be synthesized according to the methods of the present invention. While many syntheses of folic acid and related folates are available in the literature, most of these syntheses tend to  
30 be costly, multistep and tedious in nature. Thus, in accordance with preferred embodiments of the invention, the an appropriately substituted and protected pteroyl-p-aminobenzoate can be synthesized via a novel set of reactions according to the methods of the present invention. In  
35 preferred embodiments, the oxidative cleavage of folic acid is

effected using, for example, a mixture of HBr and bromine at 95-100 C. Significantly, the methods of the present invention provide an increased yield of the desired pterin-aldehyde intermediate, by performing the oxidative cleavage  
5 under conditions of elevated temperature and pressure. Thus, the present invention provides methods for the synthesis of pterin aldehyde, and methods for the preparation of folate-oligonucleotide conjugates using the pterin aldehyde.

In preferred embodiments, selective protection of the  
10 amino group of the pterin aldehyde is then performed. For example, the 2-amino group of the pterin aldehyde obtained from folic acid can be treated with isobutyric anhydride at a temperature of 130°C for 30 hours to afford N2-isobutyryl-6-formyl pterin. The 4-oxo moiety of the pterins may also be  
15 protected as the diphenylcarbamate species by using diphenylcarbamoyl chloride as reagent.

In accordance with further preferred embodiments, the N2-isobutyryl-6-formyl pterin intermediates can then be subjected to reductive amination using commercially available  
20 p-aminobenzoic acid with, for example, pyridine-borane complex in the presence of PPTS, as proton source, to produce N2 isobutyrylpterioic acid. Preferably, the intermediate is then further protected by, for example, conversion of the N10-functionality to a N<sup>10</sup>-trifluoroacetyl group via reaction  
25 with, for example, trifluoroacetic anhydride at room temperature.

In further preferred embodiments, the pterioic acid derivative is coupled with the w-aminoalkoxy nucleoside to obtain a folate-nucleoside conjugate. This coupling of the  
30 pterioic acid derivative and the w-aminoalkoxy nucleoside may be performed using any one of a variety of reagents, including but not limited to HONB/DCC, HATU, HBTU and PyBrop. More preferred is the use of PyBroP in the presence of a non-nucleophilic amine as a base to neutralize the HBr that is  
35 generated during the coupling reaction. The folate-

nucleoside conjugate so synthesized according to the methods of the invention is typically isolated in yields of about 40-50%. In some preferred embodiments, the folate-nucleoside conjugate is then subjected to a phosphitylation reaction using conventional reagents to afford the corresponding folate-nucleoside conjugate phosphoramidite in yields of about 60-70%.

Both a-folic acid conjugate phosphoramidites and g-folic acid conjugate phosphoramidites are synthesized following the methods of the present invention described herein. Use of such amidites derived from different nucleosides afford a variety of different oligonucleotide-folate conjugates of the present invention. Typically, the folate-nucleoside phosphoramidites of the present invention are used during automated oligonucleotide synthesis as solutions in dry dichloromethane, dimethyl formamide or acetonitrile using coupling times in excess of those used for the standard phosphoramidites.

Phosphoramidites may also be synthesized from the intermediate protected nucleoside conjugates such as, for example, the glutamic acid-nucleoside conjugate. These amidites may also be used according to the methods of the present invention to provide oligonucleotide conjugates.

The use of t-butyl protecting groups for pendant carboxylate groups of the glutamate portion of the conjugates of the present invention allows for facile cleavage of the t-butyl ester. However, in instances where this protecting group is difficult to remove using the conventional cleavage reagents of TFA and TFA in dichloromethane, other protecting groups such as the O-allyl group may be used.

The oligonucleotide-folate conjugates of the present invention are expected to exhibit enhanced transport into cells compared to the parent oligonucleotides. The conjugation with the folates facilitates the active transport of the conjugates of the present invention via receptor

mediated endocytosis following recognition of the folate by membrane folate binding receptors and transporters. Since folate binding receptors and transporters are widely present in a variety of cells and tissues, the oligonucleotide-folate  
5 conjugates of the present invention are anticipated to deliver higher levels of oligonucleotide into these cells and tissues.

As used herein, the term "alkyl" includes but is not limited to straight chain, branch chain, and alicyclic  
10 hydrocarbon groups. Alkyl groups of the present invention may be substituted. Representative alkyl substituents are disclosed in United States Patent No. 5,212,295, at column 12, lines 41-50, hereby incorporated by reference in its entirety.

15 As used herein, the term "aralkyl" denotes alkyl groups which bear aryl groups, for example, benzyl groups. The term "alkaryl" denotes aryl groups which bear alkyl groups, for example, methylphenyl groups. "Aryl" groups are aromatic cyclic compounds including but not limited to  
20 phenyl, naphthyl, anthracyl, phenanthryl, pyrenyl, and xylyl.

As used herein, the term "alkanoyl" has its accustomed meaning as a group of formula  $-C(=O)-alkyl$ . A preferred alkanoyl group is the acetoxy group.

25 In general, the term "hetero" denotes an atom other than carbon, preferably but not exclusively N, O, or S. Accordingly, the term "heterocycle" denotes a cyclic structure having at least one non-carbon atom.

Additional objects, advantages, and novel features of  
30 this invention will become apparent to those skilled in the art upon examination of the following examples, which are not intended to be limiting.

For the following examples, materials were obtained from Aldrich Chemical Co, RI chemicals, Cal Biochem and Sen  
35 Chemicals. The dry solvents were purchased from Aldrich

Chemical Co, and were used without further drying.  
Aminolink-2 was purchased from ABI and 3'-Amino-Modifier, 5'-  
Amino-Modifier C6 and Amino-Modifier dT reagents were  
purchased from Glen Research Corporation. All reactions were  
5 monitored by TLC (Merck Kieselgel 60 F254 plates); flash  
chromatography was performed over silica gel (Merck).  $^1\text{H}$  NMR  
was recorded at 200 or 400 MHz (Varian Unity-400 instrument).  
 $^{13}\text{C}$  NMR was recorded at 75 MHz. When modified phosphor-  
amidites were used, a longer coupling time (10-15 min) was  
10 employed. HPLC was performed on a Waters 600E instrument  
equipped with a model 991 detector. Mass spectra were  
obtained either by FAB or Electron spray method.

All oligonucleotide sequences are listed in a standard  
5' to 3' order from left to right.

## 15 EXAMPLES

### EXAMPLE 1

#### 6-Formylpterin (21)

Folic acid (20g, 45 mmole) was added to a stirred  
solution of bromine (36.2g, 226.5 mmole) in 40% hydrogen  
20 bromide (120 ml) in a one liter pressure bottle. When all  
the folic acid was dissolved, the solution was heated to 90-  
95°C for two hours with continuous stirring. The resulting  
yellow precipitate was collected by filtration, washed  
thoroughly with water and the solid cake was extracted with  
25 hot acetone. The product was dried in vacuum at 50°C to  
furnish the desired 6-formylpterin 21 as a yellow powder,  
yield = 6.7g 78% ,  $^1\text{H}$  NMR (TFA):  $\delta$  10.21 (s, 1H), 9.46 (s,  
1H). These chemical shifts agree with standard 6-  
formylpterin obtained from Aldrich Chemical Co. Mass spec,  
30 m/z,  $\text{M}-\text{H}^+ = 192$  for  $\text{C}_7\text{H}_5\text{N}_5\text{O}_2$

### EXAMPLE 2

#### $\text{N}^2$ -isobutyryl-6-formylpterin (25)

6-Formylpterin 21, (5g) was suspended in isobutyric

anhydride (250 ml) and heated under argon at 130°C with continuous stirring until all solid had dissolved (30 hours). The reaction mixture was cooled to room temperature, and then placed in refrigerator for 24 hours. The precipitated  
5 product was collected by filtration and washed thoroughly with hexanes and the last with 1:1 mixture of hexanes/IPA. The light cream colored N<sup>2</sup>-iBu-6-formylpterin was dried invacuo at 50°C, and resulted in 5g (75%) clean product, which was used without further purification. <sup>1</sup>H NMR (DMSO):  
10 10.10 (s, 1H), 9.30 (s, 1H) 2.80-2.86 (m, 1H), 1.15-1.18 (2s, 6H).

### EXAMPLE 3

#### N<sup>2</sup>-isobutyryl-pteronic acid (26)

N<sup>2</sup>-iBu-6-formylpterin (25) 4.0 gm, 15 mmole) was  
15 suspended in MeOH. To this suspension was added 4-aminobenzoic acid (2.1 g, 15 mmole) followed by stirring for 10 minutes. The reaction mixture was then cooled in an ice bath with continuous stirring. A solution of pyridinium para-toluene sulfonate (PPTS, 1 M in MeOH, 15 ml, 15 mmol)  
20 was added in one portion, followed by the addition of pyridine-borane complex (1.9 ml of 8M solution, 15 mmol) dropwise over a period of 10 minutes. The ice bath was removed, and stirring was continued at room temperature for 3 hours. The fine precipitated product was collected by  
25 centrifugation and washed with a 1:1 mixture of hexanes - MeOH several times, and finally with MeOH. The product was dried in vacuum over P<sub>2</sub>O<sub>5</sub> for 24 hours to give light cream-colored powder, yield = 5.3 gm, 91%. <sup>1</sup>H NMR (DMSO): δ 8.87 (s, 1H), 4.62 (s, 2H), 2.75-2.82 (m, 1H), 1.13, 1.16 (2s,  
30 6H); MS, m/z, M-H<sup>+</sup> = 383 for C<sub>18</sub>H<sub>18</sub>N<sub>6</sub>O<sub>4</sub>.

### EXAMPLE 4

#### N<sup>2</sup>-isobutyryl-N10-trifluoroacetylpteronic acid (27)

N<sup>2</sup>-isobutyryl-pteronic acid 26 (1.5 gm, 3.9 mmol) was

dried over  $P_2O_5$  for 24 hours under reduced pressure and then placed in a round bottom flask. To this was added trifluoroacetic anhydride (32 ml) dropwise over 30 minutes with continuous stirring. The reaction mixture was protected from light and moisture and was stirred at room temperature until all solid had dissolved, and the stirring was continued for an additional one hour. The trifluoroacetic anhydride was then removed under reduced pressure and MeOH was added to the remaining dark syrup, followed by stirring for one hour. A cream-colored precipitate was obtained which was collected by filtration and washed with MeOH three times and then twice with  $CH_2Cl_2$ . The material was dried for 24 hours over  $P_2O_5$  under reduced pressure to yield 1.26 g (70%) of compound 27 as a light cream-colored powder.  $^1H$  NMR (DMSO):  $\delta$  8.91 (s, 1H), 5.25 (s, 2H) 2.76-2.83 (m, 1H). 1.13, 1.17 (2s, 6H) MS m/z ,  $MH^+=479$  for  $C_{20}H_{17}N_6O_5F_3$ .

#### EXAMPLE 5

##### 2'-O- (N-Fmoc- $\alpha$ -tert.butyl-g-L-glutamyl )-aminohexyl- 5'-O-DMT-uridine (30)

To a solution of N-Fmoc- $\alpha$ -tert.butyl protected L-glutamic acid 28 (1.3 g, 3.0 mmol) in anhydrous DMF (7 ml) was added DCC (1.0 gm, 4.8 mmol) and HONB (0.85gm, 4.7 mmol) successively and the reaction was allowed to stir for 30 minutes to make the active ester. After this time compound 18 (2.3 gm, 3.5 mmol) was added to the reaction mixture and the stirring was continued for 7 hours, when the tlc system (50:40:10 EtOAc: hexane:MeOH) indicated the completion of the reaction. It was diluted with EtOAc (200 ml), washed with brine (3x25 ml). The organic layer was dried over anhydrous  $Na_2SO_4$ , concentrated to a foam, and purified by column chromatography on silicagel using the above solvent system to furnish a residue which had couple of impurities. The product was repurified by column chromatography using 30% acetone in  $CH_2Cl_2$ . The appropriate fractions were collected

and concentrated to furnish

30 as a white foam, yield = 2.0 gm, 61%.  $^1\text{H}$  NMR (  $\text{CDCl}_3$  ):  $\delta$  9.60 (bs, 1H), 8.00 (d, 1H,  $J = 8.2$  Hz), 7.74 (d, 2H,  $J = 7.3$  Hz), 7.61-7.57 (m, 2H), 7.37-7.25 (m, 13 H) 6.85-6.81 (m, 4H  
5 ), 6.22 (m, 1H), 5.93 (bs, 1H), 5.83 (d, 1H,  $J = 7.9$  Hz), 5.55 (d, 1H,  $J = 7.9$  Hz) 4.44 - 2.82 (m, 21 H), 2.23-1.04 (m, 19H + some DCU).

#### EXAMPLE 6

2'-O- (N-Fmoc-  $\alpha$ -tert.butyl-g-L-glutamyl )-aminohexyl-5'- O-  
10 DMT-5-methyl-uridine (31)

This compound was prepared by following the similar method as described for compound 30. Thus from 28 (2.6 gms, 6.1 mmol) 29 (4.6 gm, 6.9 mol), DCC ( 2.0 gm, 9.6 mmol) and HONB (1.7 gm, 9.5 mmol) in anhydrous DMF (15 ml) , after 10  
15 hours of stirring at room temperature and the usual workup described for compound 30, followed by the chromatographic purification using 5% isopropanol in  $\text{CH}_2\text{Cl}_2$  as the eluant to give desired conjugate as a foam, yield = 4.5 gm, 60%

#### EXAMPLE 7

20 2'-O- ( $\alpha$ -tert.butyl-g-L-glutamyl)-aminohexyl-5'-O-DMT-uridine (32)

A solution of 30 (0.76 gm, 0.72 mmol) was stirred in 0.1 M DBU (20 ml) for 1 hour at which time the tlc system (7% MeOH in  $\text{CH}_2\text{Cl}_2$ ) showed the completion of the reaction. The  
25 solvent was removed under reduced pressure and the residual oil was purified by flash chromatography on silicagel using 10% MeOH in  $\text{CH}_2\text{Cl}_2$ . The fractions containing the desired compound were concentrated to furnish the free amino acid as a white foam, yield = 0.40 gm, 67%.

#### 30 EXAMPLE 8

2'-O- ( $\alpha$ -tert-butyl-g-L-glutamyl)-aminohexyl- 5'-O-DMT-5



**methyl -uridine (33)**

This compound was essentially prepared by the same method as described for 32 with some minor changes in the workup. Therefore, compound 31 (3.8 gm, 3.56 mmol) was  
5 treated with 0.1M DBU ( 100 ml ) for 1.2 hours, at which time the deprotection was complete. The solution was concentrated to dryness and the residue dissolved in EtOAc (100 ml) and washed with water (2x 30 ml). The organic layer dried as usual and the compound purified by column  
10 chromatography using 7 % MeOH to give the desired derivative 33 as a colorless foam, yield = 2.3 gms, 76% <sup>1</sup>H NMR (DMSO-d<sub>6</sub>):  
δ 7.77-7.55 (m, 1H), 7.52 (s, 1H), 7.39-7.26 (m, 9H) 6.94-6.90 (m, 4H), 5.86 (d, 1H, J = 4.6 Hz), 5.20 (bs, exg. 1H), 4.25 (bs, 1H), 4.04-3.97 (m, 1H), 3.75 (s, 6H), 3.67-2.97 (m,  
15 6H) 2.10-2.0 (m, 3H) 1.9- 1.28 (m, 17 H)

**EXAMPLE 9****5'-O- DMT- 2'-O-aminohexyl-uridine-N<sup>10</sup>-trifluoroacetyl-α-tert-butyl-folic acid-γ-cojugate (35)**

To a solution of N<sup>10</sup>-TFA-pterioic acid, 34 from Aldrich  
20 Chemical Co. (0.1 gm, 0.24 mmol) in anhydrous DMF (1.5 ml) was added PyBrOP ( bromo-tris-pyrrolidino-phosphonium hexafluorophosphate, 0.12 gm, 0.25 mmol) and the suspension was stirred for 15 minutes. To this mixture was added DIPEA (diisopropylethylamine, 0.063 gm, 0.48 mmol) when the  
25 reaction mixture became a thick gel. After breaking the thick gel by agitation, it was stirred for 25 minutes and at this point anhydrous DMF (1ml) was added to effect a suspension. To this suspension was added nucleoside-linked glutamic acid derivative 32 (0.20 gm, 0.24 mmol), when the  
30 reaction became a clear solution in 5-7 minutes. The mixture was stirred for 15 hours after which time the tlc system (15% MeOH/CH<sub>2</sub>Cl<sub>2</sub>) showed the completion of the reaction. The reaction mixture was diluted with CH<sub>2</sub>Cl<sub>2</sub> (30 ml) and purified by column chromatography on silicagel using 10% MeOH /CH<sub>2</sub>Cl<sub>2</sub>

as the eluent. The fluorescent fraction along with the other less polar impurities was eluted fairly quickly because of the presence of DMF in the reaction mixture. The appropriate fractions were collected and concentrated to minimum volume and then addition of  $\text{CH}_3\text{CN}$  (1ml) furnished a solid which was collected by filtration and washed with  $\text{CH}_3\text{CN}$ , the solid was dried under vacuum for 3 hours to give the desired nucleoside-folic acid conjugate (**35**) as a single spot compound yield = 0.085 gm. The mother liquor was concentrated and re-chromatographed using 7-10%  $\text{MeOH}/\text{CH}_2\text{Cl}_2$  and after usual treatment furnished an additional 0.035 gm, giving final yield of 40% of the desired compound.  $^1\text{H}$  NMR, 400 MHz ( $\text{DMSO}-d_6$ ):  $\delta$  11.37 (s, 1H), 8.76 (d, 1H,  $J = 7.2$  Hz), 8.61 (s, 1H), 7.88 (d, 2H,  $J = 8.4$  Hz), 7.78 (bs, 1H), 7.72 (d, 1H,  $J = 8.0$  Hz), 7.61 (d, 2H,  $J = 8.0$  Hz), 7.37-7.22 (m, 9H), 6.88 (d, 4H,  $J = 8.8$  Hz), 5.77 (bs, 1H), 5.27 (d, 1H,  $J = 8.4$  Hz) 5.12-5.10 (m, 3H), 4.25-4.15 (m, 2H), 3.94-3.87 (m, 2H), 3.72 (s, 6H), 3.59-3.47 (m, 2H), 3.82-3.20 (m, 7H), 2.98-2.97 (m, 2H), 2.18-1.88 (m, 4H), 1.47-1.22 (m, 13 H);  $^{13}\text{C}$  NMR ( $\text{DMSO}-d_6$ ):  $\delta$  171.03, 165.61, 162.99, 158.11, 156.74, 155.77, 155.40, 154.07, 150.26, 149.39, 144.62, 144.00, 141.52, 140.21, 135.33, 135.05, 134.44, 129.75, 128.68, 128.37, 128.11, 127.90, 127.68, 126.79, 117.52, 114.66, 113.23, 101.45, 87.113, 85.87, 82.60, 80.86, 80.55, 69.81, 68.45, 62.59, 55.03, 53.78, 53.15, 31.74, 29.06, 27.63, 26.32, 26.21, 25.12; ESMS (-ive ion): M-1, 1219.3 which agrees with  $\text{C}_{61}\text{H}_{67}\text{N}_{10}\text{O}_{14}\text{F}_3$

#### EXAMPLE 10

5'-O- DMT-2'-O- aminohexyl-uridine- $\text{N}^{10}$ -trifluoroacetyl- $\alpha$ -tert-butyl-folic acid- $\gamma$ -cojugate 3'-phosphoramidite (**36**)

To a suspension of compound **35** (0.11 gm, 0.09mmol) was added DIPEA (0.029 gm, 0.22 mmol) and the reaction mixture was stirred for 10 minutes when 2-cyanoethyl N,N-diisopropylchlorophosphoramidite (0.027 gm, 0.11 mmol). The

reaction mixture was stirred for 2 hours when the reaction was complete. The reaction mixture was loaded directly onto silicagel column and it was eluted with 10% MeOH/CH<sub>2</sub>Cl<sub>2</sub>, the appropriate fractions were collected and concentrated to give  
5 the desired phosphoramidite as solid. yield = 0.085 gms, 66%;  
<sup>31</sup>P NMR (CDCl<sub>3</sub>): δ 150.91, 150.52 (with some impurity 8%, at 121.91 ppm).

**EXAMPLE 11**

2'-O-(N-Fmoc-g-tert.butyl-a-L-glutamyl)-aminohexyl- 5'-O-  
10 DMT-uridine (38)

To a solution of N-Fmoc-L-glutamic acid-g-tert.butyl ester (1.5 gm, 3.25 mmol) in anhydrous DMF (7 ml) was added DCC (1.0 gm, 4.9 mmol) and HONB (0.85 gm, 4.74 mmol) successively and the reaction mixture was stirred at room  
15 temperature for 35 minutes, when a suspension was observed. To this stirring mixture was added nucleoside (2.3 gm, 3.56 mmol) and the stirring was continued for 2.5 hours at room temperature, after which time the reaction was complete, as judged by tlc (10% MeOH/CH<sub>2</sub>Cl<sub>2</sub>). The reaction mixture was  
20 diluted with EtOAc (50 ml) and washed with water (2x20 ml). The black water layer was extracted with EtOAc (20 ml) and the combined organic layer was dried over anhydrous Na<sub>2</sub>SO<sub>4</sub> and the solvent removed under reduced pressure to give an oil which was purified by column chromatography using 5-10%  
25 MeOH/CH<sub>2</sub>Cl<sub>2</sub> as the eluant. The appropriate fractions were collected and concentrated to give an oil which after drying under vacuum for 5 hours gave a foam yield = 2.2 gm, 60%

**EXAMPLE 12**

2'-O-(N-Fmoc-g-tert.butyl-a-L-glutamyl)-aminohexyl-5'-O-DMT-  
30 5-methyl-uridine (39)

This compound was prepared by the same method as described for compound 38. Thus from 1.3 gms (3.0 mmol) of 37, 2.3 gm (3.48 mmol) of 29, 1.0 gm (4.8 mmol) of DCC and

0.8 gm (4.4 mmol) of HONB, 2.0 gm (62%) of pure product, along with additional 1.0 gm (30%) of 85% pure product of desired conjugate was obtained.

**EXAMPLE 13**

5 **2'-O-(g-tert-butyl-a-L-glutamyl)-aminohexyl-5'-O-DMT-uridine (40)**

A solution of **38** (0.5 gm, 0.47 mmol) in 0.1 M solution of DBU (15 ml) was stirred at room temperature for 1 hour. The reaction was complete as judged by tlc (10% MeOH/ CH<sub>2</sub>Cl<sub>2</sub>) analysis which was followed by the removal of the solvent under reduced pressure. The residue was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (50 ml ) and was washed with water (20 ml). The organic layer was concentrated and purified by flash chromatography on silicagel using 7% MeOH in CH<sub>2</sub>Cl<sub>2</sub> as the eluant. The fractions  
10 containing the desired compound were concentrated and dried to furnish **40** as a colorless foam , yield = 0.275 gm, 70%.

**EXAMPLE 14**

**2'-O-(g-tert.butyl-a-L-glutamyl)-aminohexyl-5'-O-DMT-5 methyl-uridine (41)**

20 A solution of **39** (2.4 gm, 2.25 mmol) was stirred with 0.1 M DBU (70 ml) for 1 hour. The reaction was judged complete (tlc, 10% MeOH/CH<sub>2</sub>Cl<sub>2</sub>) and was diluted with EtOAc (100 ml), washed with saturated sodium bicarbonate (2x30 ml), water (30 ml) and the organic layer concentrated and purified  
25 by column chromatography using 7% MeOH in CH<sub>2</sub>Cl<sub>2</sub> as the eluant. The removal of the solvent furnished the desired compound **41** as a foam, yield = 1.2 gm, 60%

**EXAMPLE 15**

30 **5'-O-DMT-2'-O-aminohexyl-uridine-N<sup>10</sup>-trifluoroacetyl-g-tert-butyl-folic acid-α-conjugate (42)**

To a suspension of N<sup>10</sup>-TFA-pterioic acid (0.094 gm, 0.23

mmol) in anhydrous DMF (1 ml) was added DIPEA (0.062 gm, 0.48 mmol) and the suspension was agitated to form a thick slurry. To this was added PyBrOP (0.115 gm, 0.24 mmol) and the stirring was continued for an additional 30 minutes. After 5 this time compound 40 was added to the reaction mixture and the clear solution was allowed to stir for 24 hours when the tlc analysis 10% MeOH in CH<sub>2</sub>Cl<sub>2</sub> indicated that some starting material was still left unreacted. The stirring was continued for an additional 12 hours, CH<sub>3</sub>CN (5 ml) was added, 10 and the reaction mixture was left at 4°C overnight. The separated solid was removed by filtration and washed with a small volume of CH<sub>3</sub>CN. The solid material was purified by column chromatography on silicagel using 10% MeOH in CH<sub>2</sub>Cl<sub>2</sub> as the eluant. The appropriate fractions were collected and 15 concentrated to furnish nucleoside-folic acid- $\alpha$ -conjugate 42 as a light cream colored solid, yield = 0.071gm, 25%, <sup>1</sup>H NMR, 400 MHz (DMSO-d<sub>6</sub>)  $\delta$  11.51 (bs, 1H), 11.39 (bs, 1H), 8.62 (s, 1H), 8.50 (d, 1H, J = 7.2 Hz), 7.88 (m, 3H), 7.78 (d, 1H, J = 8.4 Hz), 7.62 (d, 2H, J = 8.0 Hz), 7.40-7.20 (m, 9H), 6.68 20 (d, 4H, J = 8.4 Hz), 5.80 (bs, 1H), 5.30 (d, 1H, J = 8.4 Hz), 5.14-5.10 (m, 3H), 4.35- 4.25 (m, 1H), 4.21- 4.18 (m, 1H), 3.92-3.83 (m, 2H), 3.66 (s, 6H), 3.59-3.45 (m, 2H), 3.38-3.20 (m, 3H), 3.17-3.00 (m, 2H), 2.26-2.20 (m, 2H), 2.03-1.80 (m, 2 H), 1.58-1.14 (m, 15 H); <sup>13</sup>C NMR (DMSO-d<sub>6</sub>): d 171.67, 25 170.82, 165.41, 162.99, 160.68, 158.11, 156.76, 155.77, 155.43, 154.01, 150.26, 149.39, 144.64, 143.98, 141.43, 140.21, 135.33, 135.05, 134.54, 129.75, 128.55, 128.09, 127.89, 127.68, 126.79, 117.54, 114.67, 113.23, 101.45, 87.13, 85.87, 82.58, 80.86, 79.61, 69.81, 68.45, 62.59, 55.03, 30 53.80, 52.91, 31.59, 29.06, 28.98, 27.67, 26.91, 26.12, 25.10 ; ESMS (-ive ion): M-1, 1219.1 which agrees with C<sub>61</sub>H<sub>67</sub>N<sub>10</sub>O<sub>14</sub>F<sub>3</sub>

**EXAMPLE 16**

5'-O-DMT-2'-O-aminohexyl-5-methyl-uridine-N<sup>2</sup>-ibu-N<sup>10</sup>-trifluoroacetyl-a-tert-butyl-folic acid-g-cojugate (43)

To a stirring solution of N<sup>2</sup>-ibu-N<sup>10</sup>-trifluoroacetyl pterioic acid **27** (1.0 gms, 2.0 mmol, 70 % pure by NMR) in anhydrous DMF (10 ml) and to this was added PyBrOP (1.1 gm, 2.35 mmol) and the reaction mixture was stirred at room temperature for 10 minutes. After this time, DIPEA (1.1 ml, 6.3 mmol) was added and the stirring was continued for an additional 15 minutes. Then the nucleoside-glutamic acid derivative **31** (1.8 gm, 2.13 mmol) was added and the stirring was continued overnight. The tlc analysis (10% MeOH in CH<sub>2</sub>Cl<sub>2</sub>) showed that the reaction was complete. It was diluted with EtOAc (100 ml) and washed with saturated sodium bicarbonate solution (2x30 ml), water (30 ml) and brine (30 ml). The organic layer was dried and concentrated to furnish a orange colored foam, which was purified by flash chromatography using 7% MeOH/CH<sub>2</sub>Cl<sub>2</sub> as the eluant. The collection and the concentration of the appropriate fractions gave a yellow colored foam, which still showed the presence of some impurities along with the desired compound. The further purification was achieved by chromatography using 3% MeOH in CH<sub>2</sub>Cl<sub>2</sub> and concentrating the desired fractions to furnish **43** as a light yellow foam yield = 0.85 gms 31% (48% based on 70% pure pterioic acid). <sup>1</sup>H NMR 200 MHz (DMSO-d<sub>6</sub>) δ 12.40 (bs, 1H), 12.00 (bs, 1H), 11.40 (bs, 1H), 8.92 (s, 1H), 8.81 (d, 1H, J = 7.2 Hz), 7.94 (d, 2H, J = 8.3 Hz), 7.85-7.79 (m, 2H), 7.71 (d, 2H, J = 8.2 Hz), 7.52 (s, 1H), 7.43-7.25 (m, 9H), 6.94-6.89 (m, 4H), 5.86 (d, 1H, J = 4.4 Hz), 5.77 (s, 2H), 4.32-4.18 (m, 2H), 4.04 (m, 2H), 3.75 (s, 6H), 3.67-3.44 (m, 2H), 3.30 (m, 2H), 3.07-2.84 (m, 2H), 2.89-2.72 (m, 1H), 2.30-1.85 (M, 4H), 1.50-1.13 (m, 26H); high resolution FAB MS (MH<sup>+</sup>) m/e 1305.5389 (C<sub>66</sub>H<sub>76</sub> N<sub>10</sub>O<sub>15</sub>F<sub>3</sub>) requires, 1305.5444; (MNa+) 1328.00.

#### EXAMPLE 17

5'-O-DMT-2'-O-aminohexyl-5-methyl-uridine-N<sup>2</sup>-ibu-N<sup>10</sup>-trifluoroacetyl-a-tert-butyl-folic acid-g-conjugate 3'-

**phosphoramidite (44)**

To a solution of nucleoside-folic acid conjugate **43** (0.50 gm, 0.38 mmol, 90% pure) in anhydrous  $\text{CH}_2\text{Cl}_2$  (5 ml) was added DIPEA (0.2 ml, 1.14 mmol) and the solution was stirred for 5 minutes. To this was then added 2-cyanoethyl N,N-diisopropyl chlorophosphoramidite (0.136 gm, 0.57 mmol) and the reaction mixture was stirred at room temperature for 1 hour when the reaction was complete. The reaction mixture was diluted with  $\text{CH}_2\text{Cl}_2$  (50 ml), washed with saturated bicarbonate solution (2x30 ml) and brine (30 ml). The organic layer was dried over anhydrous  $\text{Na}_2\text{SO}_4$  and the solvent removed under reduced pressure. The residue was purified by column chromatography, first eluting the column with a mixture of 60:33:7 (EtOAc: hexane: MeOH) followed by 10% MeOH in  $\text{CH}_2\text{Cl}_2$ . The appropriate fractions were collected and concentrated to furnish the desired phosphoramidite as light yellow foam, yield = 0.375 gm, 66%;

$^{31}\text{P}$  NMR ( $\text{CD}_3\text{CN}$ ):  $\delta$  151.01, 150.80 ppm, high resolution FAB MS ( $\text{M}+\text{Cs}^+$ ) m/e 1637.5624 ( $\text{C}_{75}\text{H}_{92}\text{N}_{12}\text{O}_{16}\text{CsPF}_3$ ) requires 1637.5498.

**20 EXAMPLE 18**

**5'-O- DMT-2'-O-aminohexyl-5-methyl-uridine- $\text{N}^2$  ibu- $\text{N}^{10}$ -trifluoroacetyl-g-tert-butyl-folic acid-a-conjugate (45)**

This compound was prepared using the same procedure as described for compound **43**. Thus from 0.93 gms (1.94 mmol of pterioic acid analog, less than 70% pure), PyBrOP (1.08 gm, 2.31 mmol), DIPEA (1.2 ml, 6.8 mmol) and nucleoside-L-glutamic acid  $\alpha$ -conjugate **41** (1.7 gms, 2.00 mmol) in anhydrous DMF (10 ml). After the overnight reaction, and the workup and chromatography, the product **45** was obtained at 80% purity, yield = 0.40 gms, 16%, high resolution FAB MS ( $\text{MH}^+$ ) m/e 1305.5386 ( $\text{C}_{66}\text{H}_{76}\text{N}_{10}\text{O}_{15}\text{F}_3$ ) requires, 1305.5444. This compound was used in the next step without further purification.

## EXAMPLE 19

5'-O-DMT-2'-O-aminohexyl-uridine-N<sup>2</sup>-ibu-N<sup>10</sup>-trifluoroacetyl-g-tert-butyl-folic acid-a-conjugate 3'-phosphoramidite (46)

To a stirring solution of compound 45 (0.40 gm, 0.30 mmol) and DIPEA (0.123 gm, 0.95 mmol) in anhydrous CH<sub>2</sub>Cl<sub>2</sub> was added 2-cyanoethyl N,N-diisopropyl chlorophosphoramidite (0.117 gm, 0.5 mmol) slowly over 5 minutes and the reaction mixture was stirred under argon atmosphere at room temperature for 1.5 hour. The tlc system (60: 33: 7, EtOAc: hexane: MeOH) indicated the completion of the reaction. The fluorescent impurity in the starting material did not get phosphorylated. The reaction mixture was diluted with CH<sub>2</sub>Cl<sub>2</sub> (50 ml), washed with saturated sodium bicarbonate solution (2x20 ml), brine (20 ml) and the organic layer dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent removed under reduced pressure to give a residue which was purified by column chromatography on silicagel using 10-15% MeOH in a 60:30 mixture of EtOAc:hexane as the eluant. The fractions containing the product was concentrated to furnish 46 as a foam, yield = 0.343 gm, 74% (~ 25 % of the fluorescent non-phosphorus impurity was still present), <sup>31</sup>P NMR (CD<sub>3</sub>CN) δ 150.99, 150.77; high resolution FAB MS (M+Cs<sup>+</sup>) m/e 1637.5632 (C<sub>75</sub>H<sub>92</sub>N<sub>12</sub>O<sub>16</sub>CsPF<sub>3</sub>) requires 1637.5498.

## EXAMPLE 20

2'-O-(N-Fmoc-a-allyl-g-L-glutamyl)-aminohexyl- 5'-O-DMT-5-methyl-uridine (48)

To a solution of N-Fmoc protected-α-allyl-L-glutamic acid 47 (2.0 gm, 4.8 mmol) in anhydrous DMF (12 ml) was added DCC (1.51 gm, 7.31 mmol) and HONB (1.31 gm, 7.31 mmol) successively and the reaction mixture was stirred for 15 minutes, when 2'-O-hexylamino-5'-O-DMT-5-methyl-uridine 30 (3.91 gm, 5.92 mmol) was added and the reaction was allowed to stir overnight. It was diluted with EtOAc (100 ml) and



washed with water (3x20 ml), organic layer dried over anhydrous  $\text{Na}_2\text{SO}_4$  and concentrated. The residue was purified on silicagel using 3%  $\text{MeOH}/\text{CH}_2\text{Cl}_2$  as the eluant. The appropriate fractions were collected and concentrated to furnish **48** as a  
5 colorless foam, yield = 3.30 gm, 64%,  $^1\text{H}$  NMR ( $\text{DMSO}-d_6$ )  $\delta$  11.40 (bs, 1H), 7.92-7.71 (m, 5H), 7.51-7.25 (m, 13H), 6.93-6.88 (m, 4H), 5.90-5.80 (m, 2H), 5.61-5.53 (m, 1H), 5.39-5.15 (m, 3H), 4.58 (bs, 2H), 4.30-3.95 (m, 7H), 3.74 (s, 6H), 3.62-3.00 (m, 8H), 2.18-1.12 (m, 11H).

#### 10 EXAMPLE 21

##### 2'-O-(a-allyl-g-L-glutamyl)-aminoheptyl-5'-O-DMT-5 methyl-uridine (49)

A solution of compound **48** (2.0 gm, 1.9 mmol) was stirred in 10% piperidine in  $\text{CH}_3\text{CN}$  (20 ml) for 30 minutes when the tlc  
15 system (10%  $\text{MeOH}/\text{CH}_2\text{Cl}_2$ ) indicated the reaction was complete. The solvent was removed and the residue was purified on silicagel by flash chromatography method using 7%  $\text{MeOH}$  in  $\text{CH}_2\text{Cl}_2$ . The fractions were collected and concentrated to give the desired compound as a colorless foam, yield = 1.5 gms,  
20 95%.

#### EXAMPLE 22

##### 5'-O- DMT-2'-O-aminoheptyl-5-methyl-uridine- $\text{N}^2$ -ibu- $\text{N}^{10}$ -trifluoroacetyl-a-allyl-folic acid-g-conjugate (50)

A solution of **27** (0.80 gm, 1.67 mmol) in anhydrous DMF  
25 (5 ml) was stirred and to this was added DIPEA (0.50 ml, 2.90 mmol) and the stirring was continued for 5 minutes. After this time, PyBrOP (0.77 gm, 1.66 mmol) was added and the reaction mixture stirred for an additional 15 minutes followed by the addition of nucleoside-amino acid conjugate  
30 **49** (1.2 gm, 1.45 mmol) and the stirring was continued overnight at room temperature. It was diluted with  $\text{EtOAc}$  (100 ml) and then washed with saturated sodium bicarbonate

(2x20 ml) followed by brine (2x20 ml). The organic layer dried over anhydrous  $\text{Na}_2\text{SO}_4$  and concentrated. The residual solid was purified by column chromatography using 10% MeOH in EtOAc and the fractions containing the major compound were combined and concentrated. It was repurified using 5% MeOH in  $\text{CH}_2\text{Cl}_2$ . The fractions which were relatively clean were combined and concentrated to furnish compound 50 as a light yellow foam. Yield = 0.584 gm, 27% + 0.156 gm, 7% about 85% pure compound by NMR.  $^1\text{H}$  NMR (  $\text{CDCl}_3$  ):  $\delta$  12.56 (bs, 1H), 10.93 (bs, 1H), 10.20 (bs, 1H), 8.89 (s, 1H), 8.40-8.37 (m, 1H), 7.93 (d, 2H,  $J = 8.4$  Hz), 7.69 (s, 1H), 7.32-7.22 (m, 11H), 6.8 (d, 4H,  $J = 8.6$  Hz), 6.37-6.32 (m, 1H), 6.03 (d, 1H,  $J = 3.3$  Hz), 6.00-5.80 (m, 1H), 5.34-5.20 (m, 4H), 4.64-4.40 (m, 4H), 4.10-3.98 (m, 2H), 3.78-3.32 (m, 13H), 2.7-1.23 (m, 17H); ESMS (-ive ion)  $m/e$  1288.3) which is in agreement with  $\text{C}_{65}\text{H}_{70}\text{N}_{10}\text{O}_{15}\text{F}_3$

### EXAMPLE 23

5'-O- DMT-2'-O-aminohexyl-5-methyl-uridine- $\text{N}^2$ -ibu- $\text{N}^{10}$ -trifluoroacetyl-a-allyl-folic acid-g-cojugate 3'-phosphoramidite (51)

To a solution of compound 50 (0.25 gm, 0.19 mmol), DIPEA (0.057 gm, 0.44 mmol), in anhydrous  $\text{CH}_2\text{Cl}_2$  (3 ml) was added chlorophosphitylating agent (0.071 gm, 0.31 mmol) and the reaction mixture was stirred for 1.5 hours when tlc (10% MeOH) indicated the reaction was incomplete. 10% more phosphitylating reagent was added and the reaction was allowed to stir for an additional 3 hours. The reaction mixture was diluted with  $\text{CH}_2\text{Cl}_2$  (50 ml) and washed with saturated sodium bicarbonate ( 2x 20 ml ), brine ( 20 ml ) and dried over anhydrous  $\text{Na}_2\text{SO}_4$ . The solvent was removed and the compound purified by 5% MeOH in  $\text{CH}_2\text{Cl}_2$ . The fractions containing the amidite were concentrated to give a foam yield, 0.110 gm, 39%.  $^{31}\text{P}$  NMR (  $\text{CD}_3\text{CN}$  ):  $\delta$  151.01 and 150.81.

## EXAMPLE 24

## Folic Acid Derivatized Oligonucleotides

Oligonucleotide synthesis.

The oligonucleotide synthesis was performed on an ABI  
 5 380B DNA Synthesizer. A solid support bound 19-mer DMT-  
 GCATC<sub>5</sub>AG<sub>2</sub>C<sub>2</sub>AC<sub>2</sub>AT was assembled using standard reagents, cycles,  
 and procedures, phosphoramidite chemistry, and either  
 commercial oxidizer or 3H-1,2-benzodithiol-3-one 1,1-dioxide  
 (0.05 M in MeCN) as the sulfur-transfer reagent. If  
 10 introduction of **2** was planned, DMT C<sup>tBPA</sup> 2-cyanoethyl  
 phosphoramidite was used for the chain assembly.  
 Phosphoramidite building blocks 1-4 were employed on the last  
 coupling step. For their attachment, 5-methylthio-1H-  
 tetrazole was used as an activator, and the coupling time was  
 15 extended to 10 minUTES. Phosphoramidites 1-2,4 were used as  
 0.1 M solutions in MeCN. For **3**, solution in DMF (0.05 M) was  
 employed. Following the coupling of **2**, the oxidation step  
 was performed with the aid of t-butyl hydroperoxide reagent  
 (10% in MeCN; 10 min). In all cases except for **3** good  
 20 coupling yield was observed. Synthesized oligonucleotides  
 are listed in Table 1 .

Table 1

Oligonucleotides synthesized with the aid of 1-2, 4

	Amidite	Backbone	Solid support bound oligonucleotide	Deprotected oligonucleotide	ES MS, MW	
					Found	Calculated
25	1	P=S	5	8	6880.53 <sup>a</sup>	6882.72 <sup>a</sup>
	2	P=O	6	9	6218.3	6219.0
	4	P=S	7	10	6880.65 <sup>a</sup>	6882.72 <sup>a</sup>

a: tBu ester.

*Deprotection of oligonucleotides 5,7.*

30 Solid support bound oligonucleotides **5,7** were treated

with conc. aq. ammonia for 1 hour at room temperature. The solid support was removed, and the solution was kept at 50° C for 8 hours. The deprotection mixture was evaporated to dryness, dissolved in water, and separated by HPLC.

5 *Deprotection of oligonucleotide 6.*

A mixture of tetrakis(triphenylphosphine) palladium(0) (20 mg), triphenylphosphine (45 mg) and *n*-butylammonium formate (1.2 M in THF; upper phase; 1.0 mL) was shaken to obtain homogeneous solution and added to a solid support bound oligonucleotide 6. The reaction mixture was shaken at 50° C for 1.5 hours, liquid phase was withdrawn, and solid support was washed with THF (5 x 1 mL). An aqueous sodium *N,N*-diethyldithiocarbamate (0.1 M; 1.0 mL) was added, the suspension was shaken for 15 minutes, the liquid was withdrawn, and the solid support was washed with water (5 x 1 mL). Fresh portion of sodium *N,N*-diethyldithiocarbamate (0.1 M; 1.0 mL) was added, and the washing cycle was repeated. Finally, the solid support was washed with THF (3 x 1 mL) suspended in 40% aq methylamine (1.5 mL) and was kept at room temperature for 4 hours with occasional shaking. The liquid was collected, and the solid was washed with water (3 x 0.5 mL). The combined solutions were evaporated to dryness, the residue was dissolved in water (2 mL) and separated by HPLC.

*Deprotection and isolation of oligonucleotide conjugates.*

Crude oligonucleotides were applied on a DeltaPak 15 C18 300 HPLC column (3.8 x 300 mm) eluted with a linear gradient from 0 to 60% B in 60 min (0.1 M aq NH<sub>4</sub>OAc as buffer A, 80% aq MeCN as buffer B). A major peak was collected, detritylated with 80% aq AcOH for 15 minutes, re-purified on the same column, and desalted. The oligonucleotides 8-10 were characterized by ES MS (Table 1).

**EXAMPLE 25**

**Synthesis of 2'-O-[N-cyanoethoxycarbonyl-(5-aminopentyl)]-5'-O-(4,4'-dimethoxytrityl)-adenosine (53)**

To a solution of 2'-O-(5-aminopentyl)-5'-O-(4,4'-dimethoxytrityl)-adenosine (1.0 g, 1.52 mmol) in anhydrous  
5 CH<sub>3</sub>CN (8.0 mL) was added with stirring, under argon, N-(2-Cyanoethoxycarbonyl)oxy succinimide (0.43 g, 2.02 mmol) . After stirring for 5 h at room temperature the reaction mixture was diluted with EtOAc (50 mL) and washed with aq saturated bicarbonate (3x20 mL) and water (20 mL), dried  
10 (Na<sub>2</sub>SO<sub>4</sub>) and evaporated to give **53** as a light yellow foam (1.1 g, 98% yield). This compound was used without further purification for the next step.

**EXAMPLE 26**

**Synthesis of 2'-O-[N-cyanoethoxycarbonyl-(5-aminopentyl)]-5'-O-(4,4'-dimethoxytrityl)-3'-tert-butyldiphenylsilyl-adenosine (54)**

To a stirred solution of crude **53** (1.1 g, 1.5 mmol) in anhydrous DMF (5 mL) was added imidazole (0.45 g, 6.6 mmol) followed by TBDPSiCl (0.78 mL, 2.97 mmol) and the stirring  
20 was continued for 12 h. It was poured into aq saturated bicarbonate solution (25 mL) and extracted with EtOAc (2x20 mL). Organic layer separated, washed with water (2x10mL) dried over Na<sub>2</sub>SO<sub>4</sub> and evaporated to dryness. Purification by flash chromatography on silica gel (5% MeOH/CH<sub>2</sub>Cl<sub>2</sub>) gave **54**  
25 (1.13 g, 76.8%) as colorless foam. <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ 8.22 (s, 1H), 7.87 (s, 1H), 7.68 and 7.57 (2d, 4H, J=6.0 Hz), 7.45-7.18 (m, 15H), 6.74 (d, 4H, J=8.2 Hz), 6.14 (d, 1H, J=4.8 Hz), 5.79 (bs, 2H), 4.93-4.85 (m, 1H), 4.50-4.22 (m, 5H), 3.77 (s, 6H), 3.40-2.64 (m, 8H), 1.42-0.99 (m, 15H); <sup>13</sup>C NMR  
30 (CDCl<sub>3</sub>) δ 158.47, 155.44, 153.01, 148.52, 144.559, 139.520, 136.00, 135.85, 135.65, 133.19, 130.12, 129.87, 128.22, 127.79, 127.58, 126.87, 113.10, 86.63, 86.52, 84.44, 80.99, 71.83, 70.28, 63.25, 59.03, 55.25, 40.98, 40.98, 29.48,

29.11, 26.94, 23.13, 19.40 and 18.53.

#### EXAMPLE 27

Synthesis of N<sup>6</sup>-Benzoyl-2'-O-[N-cyanoethoxycarbonyl-(5-aminopentyl)]-5'-O-(4,4'-dimethoxytrityl)-3'-tert-butylidiphenylsilyl-adenosine (55)

A solution of compound 54 (0.25 g, 0.25 mmol) and benzoyl-tetrazole (0.09g, 0.50 mmol) in anhydrous CH<sub>3</sub>CN (0.5 mL) was stirred in a preheated bath at 65°C for 1.5 h. After this time, the TLC (10% CH<sub>3</sub>COCH<sub>3</sub>/CH<sub>2</sub>Cl<sub>2</sub>) that the reaction was complete. The reaction mixture was cooled to room temperature and taken up in EtOAc (30 mL), washed with aq saturated sodium bicarbonate solution (2x15 mL) and water (15 mL) dried over anhydrous Na<sub>2</sub>SO<sub>4</sub>. Solvent was removed under reduced pressure and the residue was purified by column chromatography on silica gel using 10% CH<sub>3</sub>COCH<sub>3</sub>/CH<sub>2</sub>Cl<sub>2</sub> as the eluent and the appropriate fractions were combined and concentrated to give a colorless foam, yield (0.105 g, 50%).  
<sup>1</sup>H NMR (CDCl<sub>3</sub>) δ 11.21 (s, 1H), 8.58 and 8.56 (2s, 2H), 8.04 (d, 2H, J=7.3 Hz), 7.70-7.09 (m, 22H), 6.78 (d, 4H, J=8.0 Hz), 6.23 (d, 1H, J=4.6 Hz), 4.73-4.69 (m, 1H), 4.44-4.39 (m, 1H), 4.30-4.20 (m, 1H), 4.10 (t, 2H, J=6.1 Hz), 3.71 (s, 6H), 3.25-2.79 (m, 8H), 1.39-1.02 (m, 15H).

#### EXAMPLE 28

Synthesis of 2'-O-[N-Fluorenylmethyloxycarbonyl-(5-aminopentyl)]-5'-O-(4,4'-dimethoxytrityl)-adenosine (56)

To a stirred solution of 2'-O-(aminopentyl)-5'-(4,4'-dimethoxytrityl)-adenosine (5.0 g, 7.6 mmol) in anhydrous DMF (30 mL) was added fluorenylmethyloxycarbonyl succinimide (Fmoc-Osu, 5.0 g, 14.8 mmol). The resulting solution was stirred at room temperature for 2 h. The reaction mixture was taken up in EtOAc (200 mL) washed with aq saturated bicarbonate solution (3x50 mL), water (2x50 mL) and brine

solution (50 mL). The organic layer was separated and dried over  $\text{Na}_2\text{SO}_4$ . The solvent was removed under reduced pressure to furnish a yellow foam which was purified by chromatography on silica gel, eluting the with 5%  $\text{MeOH}/\text{CH}_2\text{Cl}_2$ . The fractions  
5 containing the desired compound were combined and concentrated to afford **56** as light yellow foam (6.0 g, 90%).  
 $^1\text{H}$  NMR ( $\text{DMSO}-d_6$ )  $\delta$  8.28 (s, 1H), 8.11 (s, 1H), 7.88 (d, 2H,  $J=6.9$  Hz), 7.68 (d, 2H,  $J=7.3$  Hz), 7.44-7.22 (m, 15H), 6.84 (dd, 4H,  $J=2.4$  and 8.9 Hz), 6.03 (d, 1H,  $J=4.6$  Hz), 5.19 (d,  
10 1H,  $J=5.8$  Hz), 4.60-4.05 (m, 6H), 3.73 (s, 6H), 3.60-3.20 (m, 4H), 2.97-2.73 (m, 2H), 1.52-1.15 (m, 6H).

#### EXAMPLE 29

**Synthesis of  $N^6$ -(*N,N*-Dimethylformamidine)-2'-O-[*N*-fluorenylmethyleneoxy carbonyl-(5-aminopentyl)]-5'-O-(4,4'-  
15 dimethoxytrityl)-adenosine (**58**)**

To a solution of pyridinium *p*-toluenesulfonate (PPTS, 3.0 g, 11.9 mmol) in anhydrous DMF (8 mL) was added dimethylformamide dimethyl acetal (1.7 mL, 12.8 mmol) and the stirring was continued for 5 minutes. To this mixture was  
20 added **56** (3.0 g, 3.42 mmol) and the stirring was continued for 2h. An additional amount of DMF-acetal (1.7 mL) and PPTS (1.5 g) was added and the reaction mixture allowed stirring for 2h. It was poured into EtOAc (100mL) and washed with saturated  $\text{NaHCO}_3$  solution (2x50 mL) water (50 mL) and brine  
25 solution (50 mL). The organic layer was dried over anhydrous  $\text{Na}_2\text{SO}_4$  and the solvent removed under reduced pressure. The residue was purified by flash chromatography on silica gel using 3%  $\text{MeOH}/\text{CH}_2\text{Cl}_2$  as the eluent followed by concentrating the appropriate fractions to afford 1.6 g (52%) of **58** as  
30 light yellow foam.  $^1\text{H}$  NMR ( $\text{DMSO}-d_6$ )  $\delta$  8.91 (s, 1H), 8.39 and 8.38 (2s, 2H), 7.88 (d, 2H,  $J=6.8$  Hz), 7.68 (d, 2H,  $J=6.9$  Hz), 7.44-7.21 (m, 15H), 6.09 (d, 1H,  $J=4.2$  Hz), 5.22 (d, 1H,  $J=5.8$  Hz), 4.64-4.06 (m, 6H), 3.73 (s, 6H), 3.65-3.20 (m,

4H), 3.19 and 3.13 (2s, 6H), 1.6-1.05 (m, 6H).

### EXAMPLE 30

Synthesis of *N*<sup>6</sup>-(*N,N*-Dimethylformamidine)-2'-*O*-[*N*-fluorenylmethyloxy carbonyl-(5-aminopentyl)]-3'-*O*-*tert*-butyldiphenylsilyl-5'-*O*-(4,4'-dimethoxytrityl)-adenosine (59)

To a solution of **58** (1.1 g, 1.2 mmol) in anhydrous DMF (5 mL) was added vacuum dried imidazole (0.385 g, 5.5 mmol) and TBDPSiCl (0.74 g, 2.8 mmol) and the reaction mixture was stirred at room temperature for 3h. Ethyl acetate (100 mL) was added to the reaction mixture and was washed with aq saturated NaHCO<sub>3</sub> (3x30 mL) water (50 mL) and brine solution (30 mL). After drying the organic layer on Na<sub>2</sub>SO<sub>4</sub>, it was concentrated to dryness and purified by column chromatography on silica gel using 3% MeOH/CH<sub>2</sub>Cl<sub>2</sub> as the eluting solvent. The fractions containing the desired product were combined and concentrated to give 1.2 g (87%) of **59**. <sup>1</sup>H NMR (DMSO-d<sub>6</sub>) δ 8.88 (s, 1H), 8.35 (s, 1H), 8.26 (s, 1H), 7.88 (d, 2H, J=7.2 Hz), 7.67 (d, 4H, J=7.4 Hz), 7.58 (d, 2H, J=6.6 Hz), 7.48-7.08 (m, 19H), 6.76 (dd, 4H, J=8.8 Hz and 1.5 Hz), 6.16 (d, 1H, J=4.6 Hz), 4.75-4.69 (m, 1H), 4.43-4.20 (m, 5H), 3.71 (s, 6H), 3.5-2.80 (m, 12 H), 1.35-1.00 (m, 15H).

### EXAMPLE 31

Synthesis of 2'-*O*-[*N*-fluorenylmethyloxycarbonyl-(5-aminopentyl)]-3'-*O*-*tert*-butyldiphenylsilyl-5'-*O*-(4,4'-dimethoxytrityl)-adenosine (62)

To a solution of **56** (6.3 g, 7.2 mmol) and dried imidazole (4.2 g, 60.8 mmol) in anhydrous DMF (17 mL) was added with stirring, under argon *tert*-butyldiphenylsilyl chloride (4.2 mL, 15.0 mmol). After stirring for 3h at room temperature, the reaction mixture was diluted with EtOAc (200 mL) washed with water (4x50 mL). The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>, concentrated and the residue purified by column chromatography using 5% MeOH/CH<sub>2</sub>Cl<sub>2</sub> as the eluent. The



fractions containing the product were combined and concentrated to afford 7.0 g (87%) of **62** as colorless foam.

<sup>1</sup>H NMR (DMSO-d<sub>6</sub>) δ 8.23 (s, 1H), 8.01 (s, 1H), 7.88 (d, 2H, J=6.8 Hz), 7.72-7.56 (m, 6H), 7.48-7.09 (m, 21H), 6.77 (d, 4H, J=8.4 Hz), 6.11 (d, 1H, J=4.7 Hz), 4.75-4.65 (m, 1H), 4.43-4.15 (m, 5H), 3.72 (s, 6H), 3.30-2.80 (m, 7H), 1.35-1.00 (m, 15H).

### EXAMPLE 32

#### Synthesis of N<sup>6</sup>,N<sup>6</sup>-Dibenzoyl-2'-O-[N-

10 fluorenylmethoxycarbonyl-(5-aminopentyl)]-3'-O-tert-butyl-  
diphenylsilyl-5'-O-(4,4'-dimethoxytrityl)-adenosine (**63**)

A solution of 2'-O-[N-fluorenylmethoxycarbonyl-(5-aminopentyl)]-3'-O-tert-butyl-  
diphenylsilyl-5'-O-(4,4'-dimethoxytrityl)-adenosine **62** (6.8 g, 6.0 mmol) in anhydrous  
15 CH<sub>2</sub>Cl<sub>2</sub> (50 mL) was cooled in an ice bath to ~ -9°C and the stirring was continued for 10 minutes. To this cold solution was added under argon, N,N-diisopropylethylamine (6.5 mL, 37.6 mmol) and the reaction mixture was stirred for 5 minutes, when a solution of benzoyl chloride (2.2 mL, 18.5  
20 mmol) in anhydrous CH<sub>2</sub>Cl<sub>2</sub> (15 mL) was added slowly over 45 minutes, while maintaining the temperature of the reaction between -9 °C to -6 °C. After the addition was complete, the reaction mixture was stirred for 2.5h at this temperature followed by 30 minutes at room temperature. It was diluted  
25 with CH<sub>2</sub>Cl<sub>2</sub> (100 mL) and washed with saturated NaHCO<sub>3</sub> solution (4x50 mL) and water (2x50 mL). Organic layer was dried and concentrated to yellow foam which was subjected to purification on silica gel using 2.5% CH<sub>3</sub>COCH<sub>3</sub>/CH<sub>2</sub>Cl<sub>2</sub> as the solvent mixture for elution. Appropriate fractions were  
30 combined, concentrated to afford 6.0 g (75%) of **63** as colorless foam. <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ 8.52 (s, 1H), 8.02 (s, 1H), 7.83-7.15 (m, 37H), 6.72 (d, 4H, J=7.2 Hz), 6.18 (d, 1H, J=4.6 Hz), 4.98-4.89 (m, 1H), 4.52-4.10 (m, 6H), 3.73 (s, 6H), 3.42-2.95 (m, 6H), 1.50-1.04 (m 15H).

**EXAMPLE 33****Synthesis of N<sup>6</sup>-Benzoyl-2'-O-(5-aminopentyl)]-3'-O-tert-butylldiphenylsilyl-5'-O-(4,4'-dimethoxytrityl)-adenosine (64)**

Compound **63** (1.0 g, 0.76 mmol) was treated with 10%  
5 piperidine in anhydrous CH<sub>3</sub>CN (10 mL). After stirring the  
reaction mixture for 20 minutes, the TLC (10% MeOH/  
CH<sub>2</sub>Cl<sub>2</sub>/1.5% Et<sub>3</sub>N) indicated the disappearance of the starting  
material. The solvent was removed under reduced pressure and  
the residue azeotroped at 40°C with toluene (3x15 mL). It was  
10 diluted with CH<sub>2</sub>Cl<sub>2</sub> (50 mL), washed with water (3x20 mL) dried  
over Na<sub>2</sub>SO<sub>4</sub> and evaporated to dryness. It was purified by  
column chromatography on silica gel using a mixture of 10%  
MeOH/CH<sub>2</sub>Cl<sub>2</sub>/1.5%Et<sub>3</sub>N as the eluting solvent. Appropriate  
fractions were combined and concentrated to furnish **64** as  
15 colorless foam, yield 0.53g (70%).

**EXAMPLE 34****Synthesis of 2'-O-[(N-Fluorenylmethyloxycarbonyl-a-O-methyl)-L-glutamyl] aminopentyl-N<sup>6</sup>-benzoyl-3'-O-tert-butylldiphenylsilyl-5'-O-(4,4'-dimethoxytrityl)-adenosine (66)**

20 To a solution of N-Fmoc-a-O-methyl L-glutamic acid **65**  
(0.6 g, 1.62 mmol) in anhydrous DMF was added DCC (0.45g,  
2.18 mmol) and HONB (0.4g, 2.23 mmol). The reaction mixture  
was stirred for 5 minutes and **64** (1.4 g, 1.4 mmol) was added  
under argon in one portion. After 3h of stirring at room  
25 temperature, the reaction mixture was diluted with EtOAc (100  
mL) washed with water (3x25 mL). The organic layer was dried  
over anhydrous Na<sub>2</sub>SO<sub>4</sub> and solvent removed under reduced  
pressure. The residue was purified by column chromatography  
on silica gel using 3% MeOH/CH<sub>2</sub>Cl<sub>2</sub> as the eluent. Appropriate  
30 fractions were collected concentrated to afford 1.9 g (98%)  
of **66** (it contained about 5% contamination of DCU and was  
used for next step without further purification). <sup>1</sup>H NMR

(CDCl<sub>3</sub>) d 8.92 (s, 1H), 8.63 (s, 1H), 8.15 (s, 1H), 7.98 (d, 2H, J=6.8 Hz), 7.75-7.19 (m, 30 H), 6.75 (d, 4H, J=8.9 Hz), 6.20 (d, 1H, J= 4.8 Hz), 6.05 (bs, 1H), 5.85 (bs, 1H), 4.48-4.15 (m, 8H), 3.73 and 3.68 (2s, 9H), 3.35-3.28 (m, 2H),  
5 2.25-1.87 (m, 4H), 1.40-1.00 (m, 15 H). MS (ES) m/z 1401.4 (M+K)<sup>+</sup>.

**EXAMPLE 35**

Synthesis of 2'-O-[(a-O-methyl)-L-glutamyl]-aminopentyl-N<sup>6</sup>-benzoyl-3'-O-tert-butyldiphenylsilyl-5'-O-(4,4'-

**10 dimethoxytrityl)-adenosine (67)**

2'-O-[(N-Fluorenylmethyloxycarbonyl-a-O-methyl)-L-glutamyl]aminopentyl-N<sup>6</sup>-benzoyl-3'-O-tert-butyldiphenylsilyl-5'-O-(4,4'-dimethoxytrityl)-adenosine **66** ( 1.7 g, 1.2 mmol) was stirred with 10% piperidine in anhydrous CH<sub>3</sub>CN (20 mL) for  
15 30 minutes. The reaction mixture was poured into EtOAc (70 mL) and washed with water (5x50 mL). The organic layer was dried, concentrated and purified by column chromatography on silica gel by eluting the column with a mixture of 5% MeOH/CH<sub>2</sub>Cl<sub>2</sub>. Fractions containing the compound were combined  
20 and concentrated to yield 1.0 g (70%) of **67** as colorless foam. <sup>1</sup>H NMR (CDCl<sub>3</sub>) d 8.65 (s, 1H), 8.11 (s, 1H), 8.02 (d, 2H, J=6.7 Hz), 7.68 (d, 2H, J=6.7 Hz), 7.70-7.21 (m, 23H), 6.73 (d, 4H, J=8.9 Hz), 6.20 (d, 1H, J=5.4 Hz), 5.80-5.58 (m, 1H), 4.50-4.33 (m, 3H), 3.75 and 3.69 (2s, 10H), 3.47-3.32  
25 (m, 3H), 3.20-3.01 (m, 6H), 2.37-1.62 (m, 7H), 1.42-1.01 (m, 17H).

**EXAMPLE 36**

Synthesis of [N<sup>6</sup>-Benzoyl-3'-O-tert-butyldiphenylsilyl-5'-O-(4,4'-dimethoxytrityl) adenyl]-2'-O-(pentylamino)-N<sup>2</sup>-

**30 isobutyryl-N<sup>10</sup>-trifluoroacetyl-a-O-methyl-folic acid (68)**

To a suspension of N<sup>2</sup>-ibu-N<sup>10</sup>-TFA pteric acid (0.63 g, 1.3 mmol) in anhydrous DMF (3mL) was added under stirring

PyBrOP (0.62 g, 1.3 mmol) and the reaction mixture was stirred for 5 to 7 minutes. To this suspension was added *N,N*-diisopropyletylamine (0.69 mL, 3.9 mmol) and the stirring was continued for additional 10 minutes during which time the  
5 reaction mixture became very clear. This solution was added  
67 (0.99g, 0.62 mmol). The side of the flask was washed with DMF (2 mL). The purple/red colored solution was stirred at room temperature for 2h. It was diluted with EtOAc (50 mL) and washed with water (3x20 mL) saturated NaHCO<sub>3</sub> solution  
10 (2x20 mL) and water (20 mL). The organic layer containing the desired compound was concentrated and purified on silica gel by eluting the column with 3% MeOH/CH<sub>2</sub>Cl<sub>2</sub>. Removal of the solvent from the appropriate fractions furnished 0.88 g (63%) of 68 as light yellow foam. <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ 12.5 (bs, 1H),  
15 10.25 (bs, 1H), 9.45 (bs, 1H), 8.87 (s, 1H), 8.57 (s, 1H), 8.50 (d, 2H, J=6.4 Hz), 8.22 (s, 1H), 8.01-7.92 (m, 4H), 7.70-7.08 (m, 22H), 6.72 (d, 4H, J= 8.7 Hz), 6.35-6.31 (m, 1H), 6.18 (d, 1H, J=5.6 Hz), 5.30-5.05 (m, 2H), 4.70-4.25 (m, 3H), 3.75 and 3.72 (2s, 9H), 3.39-2.15 (m, 12 H), 1.42-1.00  
20 (m, 21H). MS (ES) m/z 1601 (M+H)<sup>+</sup>.

#### EXAMPLE 37

**Synthesis of [N<sup>6</sup>-Benzoyl-5'-O-(4,4'-dimethoxytrityl)-adenylyl]-2'-O-(pentylamino)-N<sup>2</sup>-isobutyryl-N<sup>10</sup>-trifluoroacetyl-a-O-methyl-folic acid (69)**

25 A premixed solution of triethylamine trihydrofluoride (0.50 g, 3.1 mmol) and triethylamine ( 0.16 g, 1.56 mmol) in anhydrous THF (5 mL) was added 68 (0.50 g, 0.31 mmol) and the reaction mixture was stirred at room temperature for 36h when TLC (7% MeOH/CH<sub>2</sub>Cl<sub>2</sub>) indicated only 70% reaction. An  
30 additional amount of triethylamine (0.11 mL) and triethylamine trihydrofluoride (0.28 mL) was added and the stirring continued for an additional 20h. After 60h the reaction mixture was diluted with EtOAc (50 mL) washed with aq. NaHCO<sub>3</sub> (3x20 mL) and water (20 mL). Organic layer was

dried over  $\text{Na}_2\text{SO}_4$  followed by the removal of solvent under reduced pressure. The residue was purified by column chromatography using 5%  $\text{MeOH}/\text{CH}_2\text{Cl}_2$  as the eluent. After pooling and concentration the appropriate fractions, 0.275 g  
5 (68%) of **69** was obtained as off white solid. MS (ES)  $m/z$  1361.4  $(\text{M}-\text{H})^-$ .

**EXAMPLE 38**

Synthesis of [ $\text{N}^6$ -Benzoyl-3'-O-succinyl-5'-O-(4,4'-  
dimethoxytrityl)-adenylyl]-2'-O-(pentyl amino)- $\text{N}^2$ -isobutyryl-  
10  $\text{N}^{10}$ -trifluoroacetyl-a-O-methyl-folic acid (**70**)

To a solution of **69** ( 0.18 g, 0.13 mmol) in anhydrous  $\text{ClCH}_2\text{CH}_2\text{Cl}$  (1.0 mL) was added succinic anhydride (0.017g, 0.17 mmol) and DMAP (0.007g, 0.05 mmol) and the solution stirred at 55°C (bath temperature) for 2.5h. The reaction mixture was  
15 diluted with  $\text{CH}_2\text{Cl}_2$  (30 mL) and washed with 2% citric acid solution (10 mL) followed by water (2x10 mL). The solvent was removed under reduced pressure to give colorless foam, yield 0.194 g (quantitative) and it was pure enough to be used for  
20 next reaction.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ )  $\delta$  12.5 (bs, 1H), 11.1 (bs, 1H), 9.30 (bs, 1H), 8.83(s, 1H), 8.61 (s, 1H), 8.47 (d, 1H,  $J=6.7$  Hz), 8.23 (s, 1H), 8.02 (d, 2H,  $J=8.4$  Hz), 7.90 (d, 2H,  $J=8.4$  Hz), 7.59-7.22 (m, 14H), 6.79 (d, 4H,  $J=8.7$  Hz), 6.38 (bt, 1H), 6.10 (d, 1H, 6.8 Hz), 5.54-5.51 (m, 1H), 5.32 (bs, 2H), 4.99-4.90 (m, 1H), 4.71-4.52 (m, 1H), 4.36 (bs 1H), 3.77,  
25 3.73 and 3.70 (3s, 9H), 3.60-1.03 (m, 26H).

**EXAMPLE 39****Loading 70 onto CPG**

To a solution of **70** (0.19g, 130  $\mu\text{mol}$ ) and dried DMAP (0.016g, 130  $\mu\text{mol}$ ) in anhydrous  $\text{CH}_3\text{CN}$  (1.0 mL) was added a  
30 solution of dTNP (0.045g, 143  $\mu\text{mol}$ ) in a mixture of solvent (0.7 mL  $\text{CH}_3\text{CN}/0.3\text{mL}$   $\text{ClCH}_2\text{CH}_2\text{Cl}$ ). After stirring the reaction mixture for 5 minutes at room temperature, a solution of triphenylphosphine (0.034g, 130  $\mu\text{mol}$ ) in  $\text{CH}_3\text{CN}$  (1.0 mL) was

added and the stirring was continued for 5 minutes. To this red colored solution was added activated CPG (1.0g) and the resin shaken for 2h. It was filtered washed with  $\text{CH}_2\text{Cl}_2$  and  $\text{CH}_3\text{CN}$  followed by ether until the filtrate was colorless. It was dried under vacuum and the loading was checked by trityl cation assay to be around 55mmol/g.

**EXAMPLE 40****Synthesis of 3,3-Di-O-tert-butyldiphenylsilyl-diethanolamine (72)**

10 A solution of diethanolamine **71** (2.0 g, 19 mmol) and TBDPSiCl (22.0 mL, 83 mmol) in anhydrous pyridine was stirred at room temperature for 15h. The reaction mixture was taken up in EtOAc (400 mL) washed with water (3x100 mL) and brine solution (100 mL). The aqueous layer was extracted with EtOAc  
15 (100 mL). The organic layer was combined and dried over anhydrous  $\text{Na}_2\text{SO}_4$  followed by the removal of solvent under reduced pressure to afford light yellow oil which was purified by column chromatography 1-3% MeOH/ $\text{CH}_2\text{Cl}_2$  as the eluent. The fractions containing the compound were combined  
20 and concentrated to afford colorless oil which on standing overnight solidified to a white mass. It was dried over  $\text{P}_2\text{O}_5$  for 5h to give 11.0 g (98%) of **72**.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ )  $\delta$  7.74-7.69 (m, 8H), 7.42-7.39 (m, 12H), 3.82 (t, 4H,  $J=5.2$  Hz), 2.82 (t, 4H,  $J=5.2$  Hz), 2.10 (bs, 1H), 1.08 (s, 18H).

**25 EXAMPLE 41****Synthesis of  $\text{N}^1$ -[N-fluorenylmethoxycarbonyl- $\alpha$ -O-methyl-L-glutamyl]-3,3-di-O-tert-butyldiphenylsilyl-diethanolamine (73)**

To a stirred solution of N-Fmoc-L-glutamic acid methyl  
30 ester (0.65 g, 1.7 mmol) in anhydrous DMF (5.0 mL) was added DCC (0.52 g, 2.5 mmol) and HONB (0.46 g, 2.5 mmol) and the reaction mixture was stirred for 10 minutes. To this was added **72** (1.0 g, 1.7 mmol) and the stirring was continued for

12h. After diluting the reaction mixture with EtOAc (50 mL), it was washed with saturated sodium bicarbonate solution (3x15 mL) followed by water (20 mL). The organic layer was separated and dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated to give oil.

5 This was subjected to purification on silica gel using 2.5% MeOH/CH<sub>2</sub>Cl<sub>2</sub> as the eluent. Appropriate fractions were collected and concentrated to afford 0.96 g (59%) of **73**. <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ 7.80-7.30 (m, 28H), 2.94 (d, 1H, J=7.3 Hz), 4.40-4.11 (m, 4H), 3.81-3.45 (m, 9H), 2.40-1.95 (m, 4H),

10 1.30-1.23 (m, 1H), 1.03 (s, 18H).

#### EXAMPLE 42

##### Synthesis of N<sup>1</sup>-(a-O-methyl-L-glutamyl)-3,3-di-O-tert-butyl-diphenylsilyl-diethanolamine (**74**)

A solution of **73** (0.7 g, 0.73 mmol) in 1M morpholine in

15 anhydrous CH<sub>3</sub>CN (10 mL) was stirred at room temperature for 15h. The reaction mixture was concentrated to oil, which was azeotroped with toluene (5x15 mL) to remove morpholine. The residual oil was purified by column chromatography using 5-7% MeOH/CH<sub>2</sub>Cl<sub>2</sub> as the eluent. The appropriate fractions were

20 collected and concentrated to give oil, which was homogenous on TLC and by NMR. <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ 7.64-7.58 (m, 8H), 7.44-7.34 (m, 12H), 3.79-3.37 (m, 12H), 2.45-2.38 (m, 2H), 2.15-1.95 (m, 1H), 1.80-1.60 (m, 1H), 1.39 (bs, 2H), 1.03 (s, 18H); <sup>13</sup>C (CDCl<sub>3</sub>) δ 172.38, 135.55, 133.51, 133.03, 129.90,

25 129.71, 129.47, 127.86, 127.75, 62.42, 61.75, 53.89, 51.97, 50.65, 48.26, 30.18, 29.31, 26.86, 19.13.

#### EXAMPLE 43

Synthesis of N<sup>1</sup>-[N<sup>2</sup>-isobutyryl-N<sup>10</sup>-trifluoroacetyl-a-O-methyl]folic acid-3,3-di-O-tert-butyl-diphenylsilyl-

30 diethanolamine (**75**)

To a suspension of pteric acid derivative (0.46 g, 0.96 mmol) in anhydrous DMF (1 mL) was added PyBrOP (0.45 g, 0.96 mmol) and the reaction was stirred for 5 to 7 minutes at room

temperature. To this was added DIPEA (0.25 g, 1.9 mmol) and the stirring were continued for an additional 5 minutes. This mixture was then added to the vessel containing **74** (0.5 g, 0.69 mmol). It was stirred for 2h and then diluted with EtOAc (50 mL) and washed with aq saturated NaHCO<sub>3</sub> (2x20 mL) followed by brine (15 ml). Organic layer dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated to a dark yellow foam which was subjected to column chromatography using 5% MeOH/CH<sub>2</sub>Cl<sub>2</sub> as the solvent for elution. Fractions were collected and solvent removed under reduced pressure to afford 0.554 g (68%) of **75** as a light yellow solid.

#### EXAMPLE 44

##### Synthesis of N<sup>1</sup>-[N<sup>2</sup>-ibu-N<sup>10</sup>-TFA-a-O-methyl]-folic acid conjugated-diethanolamine (**76**)

TREAHF (0.56 ml, 3.1 mmol) was dissolved in anhydrous THF (5 mL) and to this was added triethylamine (0.2 mL, 0.43 mmol) slowly over 5 minutes. The stirring was continued for additional 5 minutes and this preformed mixture was added to a RB flask containing **75** (0.4 g). The mixture was stirred for 7h. It was diluted with EtOAc (50 mL) and washed with water (3x20 mL). Organic layer was dried by usual method and purified by column chromatography with 10% MeOH/CH<sub>2</sub>Cl<sub>2</sub> as the eluting solvent. The fractions containing the compound were combined and concentrated to give 0.170 g (70%) of **76**. <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ 12.55 (bs, 1H), 10.40 (bs, 1H), 8.85 (s, 1H), 8.22 (d, 1H, J=6.8 Hz), 7.88 (d, 2H, J=8.4 Hz), 7.40 (d, 2H, J=8.4 Hz), 5.16 (bs, 2H), 4.71 (bs, 1H), 4.36 (bs, 1H), 4.18 (bs, 1H), 3.73 (s, 3H), 3.46 (s, 6H), 2.88-2.80 (m, 1H), 2.66-2.60 (m, 2H), 2.40-2.17 (m, 3H), 1.58 (bs, 1H), 1.24 (d, 6H). MS (ES) m/z 707.2 (M-H)<sup>-</sup>.

#### EXAMPLE 45

##### Synthesis of dimethoxytritylation of N<sup>1</sup>-[N<sup>2</sup>-ibu-N<sup>10</sup>-TFA-a-O-



**methyl]-folic acid conjugated-diethanolamine (77)**

A solution of compound 76 ( 0.133 g, 0.14 mmol) and dimethoxytrityl chloride (0.069 g, 0.203 mmol) in anhydrous pyridine (2.0 mL) was stirred at room temperature for 4h, at which point about 60% reaction was complete. An additional amount of DMTC1 (0.014 g) was added to the reaction mixture and stirring was continued for 2h. The reaction was poured into water (30 mL) extracted with EtOAc (3x10 mL). Solvent was removed under reduced pressure to give yellow oil which was purified by column chromatography using 3-7% MeOH/CH<sub>2</sub>Cl<sub>2</sub> as the eluent. The first fractions containing dimethoxytrityl derivative were concentrated to give light colored foam, yield, 0.040g (16%) and the major fraction containing the desired mono dimethoxytrityl compound was concentrated to afford 0.105 g, 55% of 77.

**EXAMPLE 46****Synthesis of Phosphitylation of 3-O-dimethoxytrityl-N<sup>1</sup>-[N<sup>2</sup>-ibu-N<sup>10</sup>-TFA-a-O-methyl]-folic acid conjugated-diethanolamine (78)**

To a stirring solution of 77 (0.070g, 0.07 mmol) in anhydrous CH<sub>3</sub>CN (0.5 mL) was added diisopropylamine terazolid salt (0.010g, 0.05 mmol). To this was added 2-cyanoethyl teraisopropylphosphoramidite (0.031 g, 0.102 mmol) and the stirring continued for 2h. It was diluted with EtOAc (25 mL) washed with brine (3x10 mL) and the organic layer dried over Na<sub>2</sub>SO<sub>4</sub> and solvent removed under reduced pressure to give a residue which was purified by clump chromatography to give 8 in 55% yield.

**EXAMPLE 47****30 Oligonucleotide Synthesis**

Chimeric oligonucleotides 5'-NG CAY CCC CCA GGC CAC CAT-3', 80-82, where N is a nucleotide residue derived from

phosphoramidites **79a** or **79b**, Y represents ether thymidine or [4,6-di-<sup>14</sup>C]-thymidine, and C stands for 5-methyl-2'-deoxycytidine were assembled on an ABI 380B DNA Synthesizer on 1 to 30 mmol scale using 5'-O-(4,4'-dimethoxytrityl)thymidine 3'-O-(carboxymethyloxy)acetate derivatized CPG (diglycolyl-T CPG), phosphoramidite chemistry, and either commercial iodine oxidizer or 3H-1,2-benzodithiol-3-one 1,1-dioxide (0.05 M in MeCN) as the sulfur-transfer reagent. For introduction of 5'-terminal folic acid-derivatized moiety N with the aid of phosphoramidite reagents **79a** or **79b** (0.1 M in MeCN), coupling time was extended to 10 min.

#### EXAMPLE 48

##### Preparation of <sup>14</sup>C-labeled, solid support-bound

##### oligonucleotide **81b**

[4,6-Di-<sup>14</sup>C]-5'-O-dimethoxytritylthymidine (23.2 mg, 42.6 mmol; 2.5 mCi) and 2-cyanoethyl N,N,N',N'-tetraisopropylphosphorodiamidite (16.1 mg, 53.3 mmol) in dry MeCN (130 mL) were treated with 1H-tetrazole (3.0 mg, 43 mmol) for 1 h at room temperature. The reaction mixture that contained crude 2-cyanoethyl 3'-[4,6-di-<sup>14</sup>C]-5'-O-dimethoxytritylthymidyl N,N-diisopropylphosphoramidite was diluted to 400 mL with dry MeCN and immediately used in preparation of **81b**.

A solid support-bound, 5'-deprotected oligonucleotide phosphorothioate, 5'-C<sub>5</sub>AG<sub>2</sub>C<sub>2</sub>AC<sub>2</sub>AT-3', was assembled on 20 mmol scale as described above and dried on an oil pump for 2 h. It was treated with a solution of crude 2-cyanoethyl 3'-[4,6-di-<sup>14</sup>C]-5'-O-dimethoxytritylthymidyl N,N-diisopropylphosphoramidite from the previous step and 1H-tetrazole (0.45 M in MeCN; 400 mL) for 30 min at room temperature. The solid support was washed with MeCN (10 mL) and treated with 3H-1,2-benzodithiol-3-one 1,1-dioxide (0.05

M in MeCN; 20 mL) for 10 min followed by washing with MeCN (20 mL).

The obtained solid support was replaced on the DNA synthesizer, and the chain assembly was completed as described above. Last coupling was carried out with phosphoramidite reagent **79a**. The radiolabeled oligonucleotide **81b** was deprotected with 1 M aq piperidine followed by ammonolysis, and **84b** was isolated as described below for **83b**. Crude 5'-DMTr protected oligonucleotide was dissolved in water (5 mL) and purified by semipreparative HPLC on a DeltaPak C18 column (Waters, 15 mm; 300 Å; 7.8 × 300 mm) using 0.1 M NH<sub>4</sub>OAc as buffer A, 80% aq MeCN as buffer B, and a linear gradient from 0 to 40% B in 50 min at a flow rate 4 mL min<sup>-1</sup>. Collected fractions were evaporated, the residue was treated with 80% aq AcOH for 30 min and evaporated to dryness. The obtained material was dissolved in 50% aq DMSO and loaded onto the same column. The column was washed with 0.05 M aq NaOAc (15 min) and water (15 min) at a flow rate 5 mL min<sup>-1</sup>. Elution with 60% aq MeCN and evaporation to dryness gave 28.8 mg (207 mCi; 25%) of desalted oligonucleotide **83b** (Na<sup>+</sup> salt). The material was found to be homogeneous by HPLC (Delta Pak C18; 15 mm; 3.8 × 300 mm) and capillary gel electrophoresis.

#### EXAMPLE 49

**Deprotection of folic acid-oligonucleotide conjugates 80-82 with aqueous piperidine.**

Deprotection procedure is exemplified for conjugate **80b**. After completeness of oligonucleotide synthesis, a solid support-bound **80b** (20 mmol) was treated with 1 M aq piperidine for 24 to 48 h at room temperature. The solid support was washed with another portion of the deprotecting reagent, and combined solutions were evaporated under reduced pressure. The residue was treated with ammonium hydroxide for

8 h at 55° C, and the solvent was evaporated. Crude 5'-DMTr protected oligonucleotide was dissolved in water (5 mL) and purified by semipreparative HPLC on a DeltaPak C18 column (Waters, 15 mm; 300 Å; 25 ' 100 mm) using 0.1 M NH<sub>4</sub>OAc as  
5 buffer A, 80% aq MeCN as buffer B, and a linear gradient from 0 to 40% B in 50 min at a flow rate 15 mL min<sup>-1</sup>. Collected fractions were evaporated, the residue was treated with 80% aq AcOH for 30 min and evaporated to dryness. The obtained material was dissolved in 50% aq DMSO and loaded onto the  
10 same column. The column was washed with 0.05 M aq NaOAc (15 min) and water (15 min) at a flow rate 15 mL min<sup>-1</sup>. Elution with 60% aq MeCN and evaporation to dryness gave 23.0 mg (20%) of desalted oligonucleotide **83b** (Na<sup>+</sup> salt), ESMS: found 6965.04; calculated 6966.91. The material was found to be  
15 homogeneous by HPLC (Delta Pak C18; 15 mm; 3.8 ' 300 mm) and capillary gel electrophoresis.

Analogously, deprotection of **80a** (1 mmol) followed by isolation by HPLC gave **83a** (60.3 OD, 29%), ESMS: found, 6661.23; calculated 6661.58.

20 Analogously, deprotection of **82a** (1 mmol) followed by isolation by HPLC gave **83a** (35.3 OD, 17%), ESMS: found, 6660.10; calculated 6661.58.

#### EXAMPLE 50

**Deprotection of folic acid - oligonucleotide conjugates, 80a**  
25 **and 82a, with 0.25 M sodium hydroxide.**

A solid support-bound **80a** or **82a** (1 mmol) were treated with 0.25 M aq NaOH (100 mL) for 24 h at room temperature. The reaction mixtures were carefully neutralized with acetic acid, diluted to 2 mL with concentrated ammonium hydroxide,  
30 and heated for 8 h at 55° C. After evaporation to dryness, the residue was dissolved in water, and 5'-DMTr - protected conjugates were isolated by HPLC on a Delta Pak C18 column

(Waters, 15mm, 300Å, 3.9 × 300 mm) using 0.1 M NH<sub>4</sub>OAc as buffer A, 80% aq MeCN as buffer B, and a linear gradient from 0 to 40% B in 50 min at a flow rate 1.5 mL min<sup>-1</sup>. Collected fractions were evaporated, treated with 80% aq AcOH for 15 min at room temperature, and evaporated again. Residue was dissolved in 50% aq DMSO and desalted by HPLC on the same column eluting first with 0.1 M NH<sub>4</sub>OAc for 15 min, then with water for 15 min, and finally with 60% aq MeCN to obtain the conjugate **83a** (22.8 OD, 11%). ESMS: found, 6660.72; 10 calculated 6661.58.

#### EXAMPLE 51

**Deprotection of folic acid-oligonucleotide conjugates, 80a and 82a, with concentrated ammonium hydroxide.**

A solid support-bound **80a** or **82a** (1 mmol) were treated 15 concentrated ammonium hydroxide for 8 h at 55° C. After evaporation to dryness, the desired conjugate was isolated and desalted by HPLC as described above to give **85a** (25 to 31 OD, 12 to 15%). ESMS: found, 6659.23; calculated 6660.60.

#### EXAMPLE 52

20 **Deprotection of folic acid-oligonucleotide conjugates 80a and 82a with 1 M methylamine.**

A solid support-bound **80a** or **82a** (1 mmol) were treated with 1 M aq methylamine (100 mL) for 12 h at room temperature. The reaction mixture was diluted with 25 concentrated ammonium hydroxide, and heated for 8 h at 55° C. After evaporation to dryness, isolation and desalting by HPLC gave the desired conjugate **86a** (21 to 32 OD, 10 to 15%). ESMS: found, 6673.44; calculated 6674.63.

#### EXAMPLE 53

Deprotection of folic acid-oligonucleotide conjugate 82a with tetrakis (triphenylphosphine)Pd (0) complex.

A solid support-bound 82a (1 mmol) was treated with tetrakis(triphenylphosphine)palladium(0) (57.8 mg, 50 mmol),  
5 and triphenylphosphine (131 mg, 500 mmol) , and n-butylammonium formate (1 M in THF, upper phase, pH 7.8; 1000 mL) for 2 h at room temperature. The solid support was washed with THF (3' 1 ml), and the treatment and washing was repeated. The solid support was washed with aq  
10 dimethylammonium dithiocarbamate (0.1 M; 3' 1 mL), water (3' 1 mL), and THF (3' 1 mL). The deprotection was completed by treatment with concentrated ammonium hydroxide for 8 h at 55° C. Evaporation to dryness, isolation and desalting by HPLC gave the desired conjugate 86a (14.5 OD, 7%). ESMS: found,  
15 6661.04; calculated 6661.58.

It is intended that each of the patents, applications, and printed publications mentioned in this application be hereby incorporated by reference in their entirety.

As the art-skilled will appreciate, numerous changes and  
20 modifications may be made to the preferred embodiments of the present invention without departing from the spirit of the invention. It is, therefore, intended that all such variations as fall within the true spirit and scope of the invention be encompassed by the appended claims.